UNIVERSITY OF TWENTE

GRADUATION THESIS

Determining test interval strategies for clients undergoing preventive screening

Author
Stefan GROENVELD

Supervisors
Dr. ir. Erik KOFFIJBERG
Dr. ir. Gréanne LEEFTINK
Douwe WILLINK, MSc.

Performed within the faculty of

Behavioural, Management and Social Sciences
University of Twente

February 15, 2019

UNIVERSITY OF TWENTE.

prescan

Management Summary

Preventive screening using MRI in (private) healthcare is increasingly applied. Main reasons for individuals to perform preventive screening, is to be reassured about their health status, be able to manage potential risks to avoid the fear of having regrets when not applying for screening.

Research motivation and objective Prescan is a leading organization in the provision of preventive MRI scans for individual screening. At Prescan, quality of health care services provided to their clients, is one of the main objectives they are striving for. Currently, planning a revisit to keep track of one's health state is occurring more frequently. When clients receive advice to return for screening in the future, this currently is solely based on knowledge and experiences of the involved radiologists.

To recommend an appropriate interval for a potential revisit, personalized for these specific clients, Prescan wants to have more insight in the estimation of risks in developing a clinically relevant finding. Based on data of comparable former clients and related (behavioral) risk factors, knowledge on the development of diseases can be gained. In this way, this study's core objective is the following:

"Determining test interval strategies for clients undergoing preventive screening, based on gathered and analyzed retrospective data"

Methods First, a context analysis is performed in which current procedures at Prescan are highlighted, frequently recurring definitions will be explained and relevant research is treated.

From here, the company's existing data will be gathered and prepared to visualize main trends and behavior within groups of former clients. Combined with conducting a literature search, these insights will be compared and used to see how this core problem can be approached.

By constructing a model calculating incidence rates over several subgroups of clients, three scenarios will be highlighted in the estimation of risks to develop clinically relevant findings. When plotting the incidence cumulatively over time, per subgroups risk estimations of clinically relevant findings developed per 100 person years (%) can be shown.

Results and conclusion *This section is not available in the public version.*

Recommendations When using these estimations in the determination of client advise, it is recommended to point out these indications are based on retrospective data of comparable clients. Furthermore, this study focuses solely on a selection of relevant findings, which cannot be used as generalized estimations for potential development of clinical findings directly. For further research it seems valuable for Prescan to keep track of more (behavioral) risk factors of their clients, to achieve a more individualized recommendation to revisit for preventive screening. When the amount of data increases, or retrospectively other body regions and stages of clinically relevant findings are included in the analysis, it seems possible to indicate personalized predictions of developing potential threats in future.

Management Samenvatting

Preventieve sceening door middel van MRI in de (private) gezondheidszorg wordt steeds vaker toegepast. Voornaamste redenen voor individuen om te kiezen voor preventieve screening, is om gerustgesteld te worden over hun gezondheidstoestand, om potentiële risico's te kunnen beheersen en de angst om spijt te krijgen wanneer men niet voor screening zou kiezen.

Onderzoeksmotivatie en doel Prescan is een toonaangevende organisatie in het aanbieden van preventieve MRI-scans voor individuele screening. Bij Prescan is de kwaliteit van de dienstverlening aan hun cliënten een van de belangrijkste doelstellingen die zij nastreven. Momenteel komt het plannen van een terugkeer voor een update van de gezondheidstoestand steeds vaker voor. Wanneer cliënten advies krijgen om in de toekomst voor herhaalde screening terug te keren, is dit met de huidige werkwijze uitsluitend gebaseerd op kennis en ervaringen van de betrokken radiologen.

Om een geschikt interval aan te bevelen voor een mogelijk herhaalde screening, gepersonaliseerd voor deze specifieke cliënten, wil Prescan meer inzicht krijgen in het schatten van risico's voor het ontwikkelen van een klinisch relevante bevinding. Kennis over de ontwikkeling van ziekten kan worden verkregen op basis van gegevens van vergelijkbare voormalige cliënten en gerelateerde (gedrags-) risicofactoren. Op deze manier is het belangrijkste doel van deze studie als volgt:

> *"Bepaling van screeningsinterval strategieën voor cliënten die preventieve screening ondergaan, gebaseerd op verzamelde en geanalyseerde retrospectieve gegevens"*

Methoden Eerst is er een contextanalyse uitgevoerd waarin de huidige procedures bij Prescan worden behandeld, vaak terugkerende definities worden uitgelegd en relevant onderzoek wordt besproken.

Vanaf hier werden bestaande gegevens van het bedrijf verzameld en voorbereid om belangrijke trends en gedragingen binnen groepen van voormalige cliënten te visualiseren. Gecombineerd met het uitvoeren van een literatuuronderzoek, worden deze inzichten vergeleken en gebruikt om te zien hoe dit kernprobleem kan worden benaderd.

Met het opstellen van een model voor het berekenen van incidentiecijfers over verschillende subgroepen van cliënten, zijn drie scenario's beschreven om tot een geschat risico te komen voor het ontwikkelen van klinisch relevant bevindingen. Bij

het berekenen van de cumulatieve incidentie geplot over de tijd, kunnen per subgroepen risicoschattingen van klinisch relevante bevindingen die zijn ontwikkeld per 100 persoonsjaren (%) worden aangetoond.

Resultaten en conclusie *This section is not available in the public version.*

Aanbevelingen Wanneer deze schattingen worden gebruikt bij het bepalen van het advies voor cliënten, wordt aanbevolen om te benadrukken dat deze indicaties gebaseerd zijn op retrospectieve gegevens van vergelijkbare cliënten. Bovendien is deze studie uitsluitend gericht op een selectie van relevante bevindingen, die niet direct kunnen worden gebruikt als algemene schattingen voor de mogelijke ontwikkeling van klinische bevindingen. Voor verder onderzoek lijkt het waardevol voor Prescan om meer (gedrags- en) risicofactoren van haar cliënten bij te houden, om een meer geïndividualiseerde aanbeveling te krijgen voor preventieve screening. Wanneer de hoeveelheid gegevens toeneemt, of met terugwerkende kracht andere lichaamsgebieden en stadia van klinisch relevante bevindingen worden opgenomen in de analyse, lijkt het mogelijk om gepersonaliseerde voorspellingen te geven voor het ontwikkelen van potentiële gezondheidsrisico's in de toekomst.

Contents

Acknowledgements

Well, here we are: Right at the first pages of my master thesis, but simultaneously at the final stage of my study period.

After passing my bachelor program in Health Sciences, I chose to apply for a (pre)master program in a specialized direction of Industrial Engineering and Management, called Healthcare Technology and Management. I told myself I did this with the intention to challenge myself by combining and broadening my healthcare interest with this new field of knowledge. Perhaps, it was just a decision to expand my study career with some extra years of learning and fun. But no matter what the exact reason for this decision was, it currently feels like it was a good one.

I always asked myself where this so-called healthcare passion was coming from. But, with graduating at Prescan, and the earlier internship I did at the hospital on the beautiful island of Aruba, I experienced that helping other people in one way or another makes me feel good. And what I think is most important about this feeling, is that it only works in both directions. Therefore, I'm grateful to everyone who was helping me to achieve this.

First of all, I want to thank Erik Koffijberg for the informative brainstorms we had on how to work out multiple challenges during this thesis, and Gréanne Leeftink for giving me feedback with a useful perspective and always offering me help when needed. I also want to thank Douwe Willink, for guiding me through the great company of Prescan, and always providing me with good insights. It felt good to experience how companies like this function, and I got to know many nice and interested people over there.

Of course, a special thank you to all my friends from my high school, study, sports association and NovelT for always having fun; the best energy source there is if I'd say. But most of all, I want to thank my family and especially my parents and Lonne, for their unconditional love and motivation, during my frequent and unjustified complaining about how hard the life of a graduate can be.

Stefan Groenveld

Chapter 1

Introduction

Chapter 1 provides a general introduction into the topic of preventive screening. In addition, it gives background information about Prescan, the company where the research is conducted, and the motivation of this research. In this way there will be worked towards a main definition of the research objective and a method to eventually come to a solution of the core problem.

1.1 Research motivation

In order to describe the underlying essence and motivation of the performed research, necessary information will be highlighted first.

Public and private healthcare

The Dutch (public) healthcare system In the Netherlands the healthcare system is regulated in a way that it is accessible for everyone to receive the healthcare needed. In change for this accessibility every Dutch citizen is obliged to monthly pay the so-called health insurance. To prevent this system from performing unnecessary expensive treatments, consults and other kinds of healthcare provision, it is broadly divided into the primary and secondary care. From this perspective, the general practitioners (GPs) within the first layer of this division fulfill an important gatekeeper role in the assessment of consulting individuals, and whether or not providing them with an official reference to be further treated in the secondary healthcare (hospitals). Since these GPs are professionally educated and their decisions are a result of knowledge and empirical evidence, it can be said that this type of healthcare is focusing on treating and curing patients already having a certain degree of an illness or a disease.

Focusing on the prevention and early diagnosis of disease development, this preventive screening starts playing a role. From government's health economic perspective, preventing frequently occurred chronic diseases from developing is highly relevant. Therefore, they started providing public screening programs for a substantiated selection of these diseases.

Preventive Medical Research From government's perspective and thus with regard to the greater community, screening programs they offer are to prevent people from developing a (chronic) disease. In addition, while using the strategy of a cyclic screening program, the goal of this public screening is to detect and diagnose to better cure or treat a certain disease in an early stage of development. Furthermore, an increase in frequency improves the sensitivity and specificity of the related outcomes as well. With these programs public health will be monitored more carefully, which can be profitable for future health provision; Eventually, when having the possibility to treat diseased individuals in an earlier stage of development, the amount of treating chronic diseases or other intensive cases can be reduced [1]. Aside from the screening of populations, individuals can also check or screen their own health status without having any complaints or reason for indications. Within the Dutch healthcare sector this preventive medical research is called "Preventief Medisch Onderzoek" (PMO) [2]. From an individual's perspective PMO can be performed in several ways, where the main reasons to do so is to be reassured about the status of their health, being able to manage potential risks, and avoiding the fear of having regrets later on [3].

Prescan

Prescan, a company originated in the eastern of the Netherlands, is a leading organization in this provision of private preventive screening techniques. Where most of the managerial and administrative departments are established in the headquarter of Hengelo, Overijssel, many of its private clinics are to be found just across the borders with Germany: Bottrop, Düsseldorf, Gronau, Rheine, Oberhausen and Ochtrup [4]. Clients can also go to the relatively new private centres in Baarn, Schiedam, Amsterdam and Den Bosch [5].

Besides the fact that Prescan offers multiple targeted health checks and scans, the most relevant program they offer is called the Total Bodyscan (TBS). The TBS is an extended version of a preventive MRI investigation, which includes several MRI-scans and cardiological and additional tests.

Estimating the risk to develop a disease After undergoing the first screening, clients receive feedback on their results. By doing so, Prescan wants to optimize the quality of health care services provided to their clients. Currently, planning a revisit to keep track of one's health state is occurring more frequently. To recommend an appropriate interval personalized for these specific clients, Prescan wants to have more insight in the repeating behavior of its clients and risk estimations in developing a clinically relevant finding.

In conclusion, Prescan provides their clients with a form of PMO in which bodyscans are performed. By doing so, every individual in essence receives a kind of update about their health status. An insight is given in several relevant body regions, in

which a selection of diseases and abnormalities can be diagnosed. Although in general it is unknown whether an individual will develop a disease, Prescan prefers to improve their advice for a potential revisit in future. Based on personalized risk profiles, knowledge in the development of diseases to potentially develop in future can be examined.

1.2 Problem description

In this paragraph the main objective will be concisely highlighted and described as a result of the general introduction and the problems Prescan is facing, concerning the topic of cyclic preventive screening.

Aim of the research

One of the main services within the possibilities of preventive screening techniques Prescan offers, is the Total Bodyscan (TBS), which is considered to be a set of MRI scans for different regions of the human body. To optimize the service towards existing and incoming TBS clients and thus provide better information and consultation, one of the objectives considers to scientifically proof and argue related risk factors and growth characteristics of several abnormalities and diseases. In this way, the goal is to come to a cyclic preventive screening strategy, with a personalized strategy for an incoming individual for revisiting Prescan. Based on retrospective data, some useful insights of the existing Prescan population will be given, to get to a more personalized strategy to revisit for screening for clients who underwent a TBS, without being diagnosed with an abnormality.

Currently, PMO is one of the hot topics within political discussions because of the related positive and negative consequences and experiences. Hence, this gives extra motivation to improve the reasoning and service towards Prescan's clients.

Core problem and research scope

In the current way of Prescan's working procedures, this type of PMO is provided ad hoc, where every individual undergoing a TBS is scanned request based. When applying for an MRI, Prescan will have contact with a client to plan a moment for investigation. A certain period of time after this moment, these clients are contacted for a potential repetition, or revisit. The interval frequency used, can be defined as the usual time horizon between a client visiting the screening clinic and the moment Prescan contacts again for a revisit. The determination of this interval is based on experience of physicians after analyzing imaging results, combined with the client's interest and their wish to revisit or not. Mostly, physicians' knowledge and experiences can help in this advice for a revisit. But to do so effectively, Prescan would like to gain more knowledge in development of diseases and trends in their existing client's behavior.

Therefore, the main goal of this research is to better specify and explain certain cyclic screening strategies and optimize the client's experience and interval of their Prescan visits. Especially, when no findings are diagnosed during the first screening, it may be valuable to revisit later, in terms of repeatedly performing preventive MRI to update on their health status through this medical checkup. This can be a revisit for a TBS, or a selection of (one of the) focus areas Prescan offers.

To achieve more knowledge on this, the report will be conducted with a focus on the TBS. This is because it covers the base of the screening strategies developed and is applied the majority of clients. Due to the limitation of the time window and availability of data required, the MRI of the abdomen is given priority of focus. Hence, in order to come to a general setup of achieving relevant insights to this problem, this (abdominal) part of the Total Bodyscan will be used as the starting point.

In conclusion, the core problem can be defined as followed:

"How can the time interval to a next potential screening moment be determined, for Prescan's clients undergoing a Total Bodyscan, based on gathered and analyzed retrospective data?"

Research questions

Following from the definition of the core problem and provision of background information, several research (sub)problems have been developed in order to come to a solution for the main research question:

1. What are current procedures and definitions within the topic of (private) preventive screening?
2. What are current strategies and procedures used within Prescan's organization for treating their clients?
3. What are possible outcomes and clinically relevant findings of the concerned abdominal MRI within the TBS concept??
4. Is it feasible to accurately determine the probability of developing and growing a clinically relevant finding based on currently available literature and data, and if so, can this be translated towards a personalized interval?
5. How can the steps taken to solve the core problem of this research be improved to allow generalization to the other diagnostic tests offered by Prescan?

With the above mentioned research (sub)problems, relevant questions needed to solve the core objective are presented in a logical sequence. The setup of these questions acts as an approximate outline of the report as well. Next paragraph presents the corresponding methodology used to answer this study's main objective.

1.3 Methodology

In order to achieve an actual answer to the main research question, this section is meant to explain the strategy used for giving answers to the research (sub)problems incrementally. The questions presented on the previous page serve as a guideline throughout the research.

Problem approach

To determine a solution method for the core problem, a context investigation of preventive screening will be performed. To do so, current developments are highlighted. Relevant publications and background literature will be taken into account. In this way, this functions as required knowledge and preparation to develop further towards the report's objective. In addition to the context of the topic, an internal analysis of Prescan's organization procedures will be performed. By doing so, useful insights of responsibilities and data processing will be identified.

Since the objective is to improve recommendations for future follow-up towards the clients, it is important to know what experienced (dis)advantages are within preventive screening. Psychosocial factors and preferences can contribute to shared decision making of the screening frequency. Therefore, relevant articles will be included within the investigation of the topic, before further focus of the report is determined.

After performing this context analysis where relevant insights and perspectives are treated, the company's available data will be gathered. When gathered and analyzed, all included data will be prepared in a way that the intervals of all retrospective clients can be presented. Since a lot of clients perform preventive screening with multiple possible outcomes, a certain delimitation will take place. Next paragraph will shortly treat several considerations made within this strategy.

Subsequently, a literature search will be conducted focusing on current status of incidence, golden standards and other disease-related information. With a focus on screening strategies, relevant information can be used on the database conducted within the case of Prescan.

When established all of the previous steps, the intention of this study is to eventually generalize the setup and model for overall preventive MRI findings, and thus for clients performing preventive screening in general. The execution of this step will be, logically, completely depending on the results following from all these steps and the extent to which the subproblems could have been answered.

Research scope and assumptions

As mentioned in the description of the research scope, the setup will face certain delimitations. At this moment, Prescan wants to improve scientific recommendations

for future screening provided to their clients undergoing the most used screening procedure within Prescan: the Total Bodyscan (TBS).

First, this report initially focuses on the abdominal MRI, due to the biggest potential share of possible findings and related consequences. Considering the restricted time window and amount of different screening outcomes, the possible abdominal scenarios will be demarcated. This is executed based on their prevalence, relevance and stakeholders' preferences. In order to achieve this consensus, relevant data will be analyzed first, combined with an expert elicitation to determine further focus. Afterwards, this same delimited set of focus will be applied to the literature search.

By applying this delimited set of events, other possible events and related consequences will be excluded and not taken into account which could influence client's recommended screening strategy.

According to the literature searches, a relevant part of the research is delimited by the availability of the required knowledge as well. This can be a potential threat to the research simultaneously, due to the lack of scientific substantiation. Therefore, within the determination of the preferences of potential health outcomes, the number of known screening moments and tangibility of results will be taken into account.

The next step and challenge after the selection and gathering of relevant data is to translate this into the probability of developing certain abnormalities. Although it is to be expected that the available knowledge will give a predictive value in possible future health states for individuals, the uncertainty that comes with it must be taken into account as well.

1.4 Thesis outline

In order to introduce the research topic, Chapter 1 provides background information and a brief explanation of the research questions conducted to eventually come to a solution for the core problem.

Following from this introduction, Chapter 2 gives an insight in relevant context of the problem in order to achieve a better understanding concerning preventive screening. Additionally, this part presents current methods used within Prescan's organization with a flowchart, identifies experienced pros and cons of this very recent topic and discusses some relevant and recent literature.

Chapter 3 includes the methodology part, where the conducted methods are being highlighted for gathering and analyzing all required information. Steps taken within the analysis of empirical data, and the preparation towards all single en repeatedly visiting clients will be described incrementally. Furthermore, visualizations of all clients and related screening moments are provided for some relevant insights in cyclic preventive screening.

In Chapter 4, a literature searches are conducted, to come up with studies and insights towards risk factors and probabilities of abnormality development, and related follow-up strategies.

Based on the mentioned methods, relevant information is presented in Chapter 5. Gathered data will be analyzed and discussed to come to an optimal use of the preferred directions. To strengthen the substantiation of selected findings, a clustering and comparison with existing literature will take place.

Based on the relation between the existing literature and the analyzed data, Chapter 6 describes the strategy used to come to a risk estimation for potential disease development in future, based on information from several client groups. In this chapter, the setup of all steps within the model is explained.

In Chapter 7 the output results from the model are presented, where it provides insights for individual risk estimations in different age categories.

Chapter 8 provides an overview of relevant conclusions and a recommendation.

The final Chapter 9 briefly treats the scope and delimitations of this study conducted, implications for theory and ends with recommendations for further research.

Certain parts or chapters are not included in final publication, as a result of competition sensitive information for the company considered.

Chapter 2

Context description

In order to understand the topic and to be well-informed, this chapter provides essential and additional knowledge regarding preventive screening developments and current methods within Prescan's healthcare services. In the upcoming sections the first two research questions will be taken into account.

2.1 Preventive screening

As briefly described in the introduction chapter, public screening programs and additional preventive medical research (PMO) techniques recently are topics being frequently discussed Where the last decades a paradigm shift takes place from reactive towards a more proactive attitude, patients are becoming more autonomous and curious about their own health state. Together with this behavioral trend, the availability of better developed screening techniques increases. State-of-the-art imaging techniques became established research tools and made it possible to give an update on people's health states [6]. Combined with the fact that the provision of healthcare is commercializing, it all facilitates clients to receive a medical check-up of their health state more easily [3].

Recent developments show that the interest of performing PMO is increasing and consumers are more frequently willing to pay for this type of research. Amongst existing Dutch healthcare insurance companies, ONVZ is the first but only one reimbursing this preventive type of healthcare (partly) nowadays, within their additional insurance [7][8]. In 2016, the Dutch Minister of VWS even pleaded for a higher flexibility of offering PMO, so that clients receive more protection due to the freedom of having a choice [9]. However, the Dutch doctor's federation KNMG, focusing on ensuring responsible provision of healthcare, thinks possible risks of expensive follow-up need more attention. Therefore, they have developed a guideline to ensure the quality of PMO [10].

Diagnostic versus preventive screening

Normally, patients being diagnosed with certain diseases or risks to further develop a worse health status, can be referred to perform an MRI screening. With this technique a specific scan will be applied to the patient to produce detailed images of

the specific areas of the body to be investigated. Shortly said, the MRI will be performed based on a medical indication. On the other hand, when looking at preventive screening within a general population, diagnosis of abnormalities to be treated is less common and can be therefore called incidental Based on experience from contacted radiologists at the Prescan clinic within the Mathias Spital in Rheine, it is assumed that these differences in characteristics of the population do not directly influence sensitivity or specificity of the performed scans [11]. Before we discuss these topics and related definitions in more detail, an explanation is given of what an MRI exactly is, and how Prescan uses this technique as part of their health services.

2.2 Magnetic Resonance Imaging

Technical functionalities

Magnetic resonance imaging (MRI) examinations have been confined to regions of the body covered within the field of an imager, the machine performing MRI scans. This includes imaging of an individual organ or a single body region. When it comes to whole-body screening, it may require evaluation of the entire body volume or vascular system. To perform this imaging effectively, serial acquisition of individual body regions was achieved by development of multistation and table movement techniques [12].

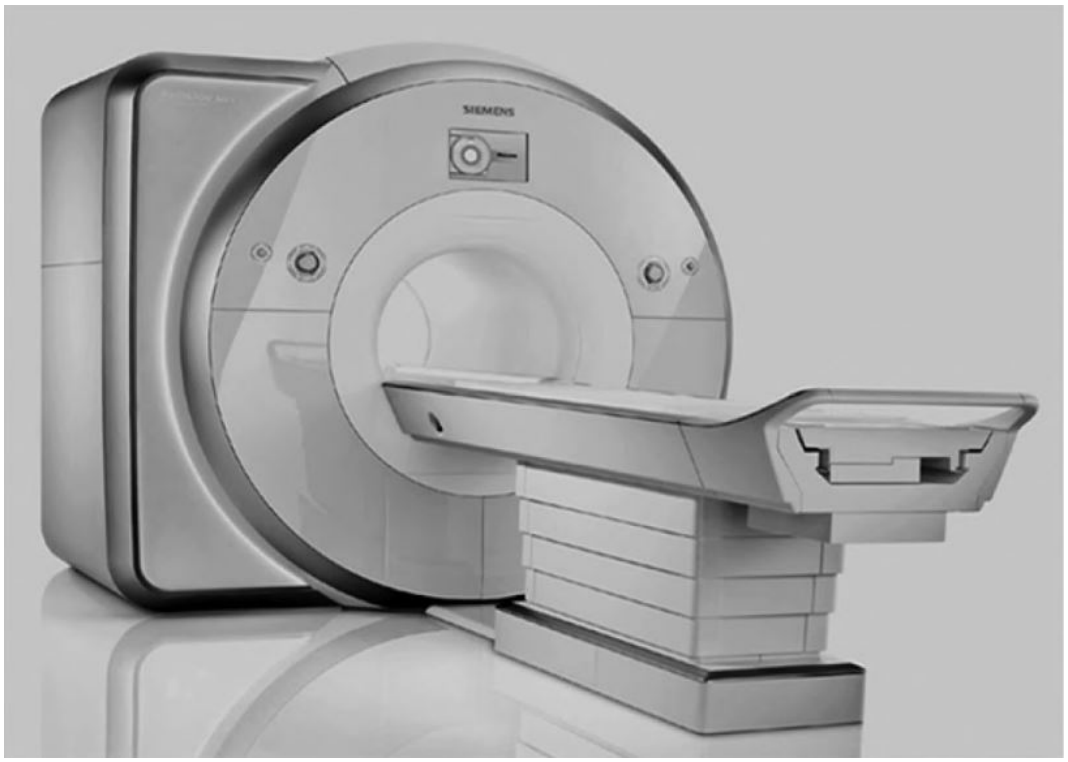


FIGURE 2.1: Example of a modern cylindrical 1.5-Tesla MR imager from Siemens, also used at some of the Prescan clinics [12]

MRI is a technique used within the radiology department of healthcare clinics, in which the use of a strong magnetic field, radiofrequency system and computer are combined to enable picturing of the internal human body. Where the main magnet of the imager must provide enough magnetization, the cylindrical configuration of it is currently the most frequently used magnet shape in medical MR systems.

With this technique, noninvasive medical tests can be performed due to the production of very detailed picture of organs, soft tissues, bones and other (soft) body structures [13].

Advantages of an MRI are, compared to other imaging techniques, that it can accurately identify abnormalities or diseases by producing images of soft-tissue structures within the body. Organs obscured by bones for example can be investigated easily with the use of MRI. Because of the 3D structure in which the body is scanned, any plane of the human body can be imaged. Besides, this imaging technique is noninvasive for the patients, and no contrast material is required. On the other hand, the amount of costs, duration and isolated experience for patients when performing a scan must be taken into account [13], [14]. Aside from the relatively low disutility of an MRI, some studies even conclude a higher sensitivity and specificity when using an MRI in asymptomatic populations [15].

Unavailable section

This section is not available in the public version.

2.3 Psychosocial factors and clinical relevance

An important and remarkable recommendation of the previously described guideline of KNMG about PMO, is that providers of PMO need a proof that profits for clients applying for preventive screening outweigh the possible risks and disadvantages. Because clients themselves have the right of self-determination, their opinion in which screening strategy to choose for is of course worth considering.

Unavailable section

This section is not available in the public version.

Unavailable section

This section is not available in the public version.

Clinical relevance

MRI screening recently gained more attention in the professional community. Initially, it was primarily accepted by potential 'clients' showing trust in these techniques, for diagnosing all possible diseases with one technique. An objective of

this screening technique should be a disease or finding which is detectable by MR. Examples of malignant diseases detectable by this whole-body MRI are bronchial carcinoma, renal carcinoma, colonic cancer and lymphoma [16].

Obviously, there is no consensus on the definition of 'incidental' or 'relevant' , which makes it important to know which perspective is used in this report. In regular healthcare, findings that are unrelated to the clinical indication for the imaging examination performed, are so-called incidentalomas. Such an incidental finding is defined as "an incidentally discovered mass or lesion, detected by CT or MRI, or other imaging modality performed for an unrelated reason" [17].

Management of incidental findings occurred in epidemiologic studies for example, is a well-recognized problem or challenge in medical research. Discussing these findings became more relevant, since frequencies increased due to the use of improved imaging techniques as research tools. Of human subjects performing screening of the brain, it is estimated that 1-8% have clinically relevant findings. Besides, recent researches suggest that in other organs examined, incidental findings are even more common [6].

Since clients come to Prescan in general to check for their health status asymptomatically, abnormalities identified through MRI researches can be considered as incidentally diagnosed as well. An even more fitting definition in this scenario, comes from Wolf's article, stated as:

"A finding concerning an individual research participant that has potential health or reproductive importance and is discovered in the course of conducting research but is beyond the aims of the study. Incidental findings, are unintended (not what they were looking for), not unexpected" [18]

Abnormalities diagnosed through Prescan's MRI examination can differ from potential health importance and life threatening, to marginal or no significance. Therefore, these findings diagnosed, incidental or not, are the ones of which the educated personnel within Prescan determines that follow-up research is needed. Therefore, we come to the following definition of this report's use of a clinically relevant finding:

"An abnormality detected by preventive MR screening of which follow-up investigation is advised to reduce risk of complications"

In Chapter 3 we will use this information as a base for further research.

Concerning the abdominal MRI, a lot of potential outcomes can be confirmed as abnormal finding of a performed preventive scan. Examples of possible abnormal findings can be different types of cancer or benign tumor growths, abdominal aortic aneurysms, obstructed veins or cyst presences in several organs [19][14][20].

Unavailable section

This section is not available in the public version.

Unavailable section

This section is not available in the public version.

2.4 Conclusions

In this chapter, relevant definitions are discussed to get prerequisite knowledge for a better understanding of this report's topic. By using insights from relevant studies performed, consequences and relevance of undergoing preventive screening are highlighted. With this context analysis and the addition of current procedures at Prescan, the first two research questions have been answered:

1. What are current procedures and definitions within the topic of (private) preventive screening?
2. What are current strategies and procedures used within Prescan's organization for treating their clients?

Chapter 3

Data collection and analysis

This chapter is not available in the public version.

Chapter 4

Literature study

This chapter, together with Chapter 5, will give insights in answering the first part of research question 4: *Is it feasible to accurately determine the probability of developing and growing a clinically relevant finding based on currently available literature and data?* In order to do so, literature will be used to gather information and see what current trends and knowledge are. In this way, insights and strategies will be used to relate to Prescan's situation and data.

4.1 Review question

Performing a literature study requires a clear question to review. Important in the definition of such a question is the specification of different parts in it. The type of population, intervention and outcome are of interest [21]. Key factors and purpose of this literature study is bridging the gap of knowledge between the analysis of the observed data and findings within literature. By comparing these two, insights within the data considered can be confirmed or strengthened. On the contrary, most probably there will not be an overload of information on studies focusing on the interval of preventive screening strategies. This is because few cases focusing on incidental screening are currently proven to be cost-effective.

4.2 Search strategy: SPICE

Priority in the formulation of a structured literature study, is to divide the stated review question into the criteria of the different SPICE-levels. The use of the SPICE framework is a method which can be used to setup structured reviews, especially when it considers a healthcare intervention [22]. Relevant search terms are gathered for the different interests, focusing on the three findings determined in the previous chapter. Therefore, the question to be answered with the use of literature will form the base, whereas the incidental finding and thus outcome of the MRI specifies this direction:

1. Aorta (aneurysm, dilatation, ectasia)
2. Pancreas (cysts, tumors)

Setting	Preventive Medical Research (PMR)
Population/Perspective	Client
Intervention	Screening
Comparison	Population without screening intervention
Evaluation	(Research) outcome: Abdominal aorta, Pancreas, Kidney

TABLE 4.1: SPICE criteria set up for literature search

3. Kidney (cysts, tumors)

By determining these specifications in order to retrieve information about strategies for preventive screening, the following review question is defined:

"From the perspective of the selected abnormalities, do there exist standards or models for development, progression, and related screening follow-up?"

Related SPICE criteria and several related terms are defined in Tables 4.1 and 4.2 to cover all important search terms to get to the base of relevant articles, based on the information to be gathered.

Now that the different criteria have been established for this specific literature question, related search terms will be set up. Per level within the constructs, multiple search terms are determined. Initially, this is done by mentioning related terms per level. Afterwards, to improve search queries, these terms are both broadened and narrowed. In this way, based on the quality of results, search strings can be adjusted with or without adding some of these.

The next step in a sound literature study is to determine the database or search engine in which a search will be performed to find the literature required in answering the research question. Initially, the database used in this literature study is determined to be PubMed, due to the enormous amount of relevant references to journal articles with a concentration on biomedicine.

After performing this literature search, the content and sufficiency will be evaluated, to conclude whether solely using PubMed would be enough.

Article inclusion criteria

Prior to the use of different search term combinations, additional selection criteria will be determined to sharpen and improve the literature study. By the construction of several restrictions, some irrelevant, outdated, or unavailable articles can be automatically excluded. Furthermore, the aim is to provide relevant and understandable information to answer the research question.

First, because preventive medical research is recently increasing in popularity, articles published in the previous millennium are considered to be outdated. Therefore, articles published after 2000 are included. A priori defined inclusion criteria were as followed:

Constructs	Related terms	Broader terms	Narrower terms
X1: Research outcome	Finding Result (optimal) Interval	"Health state"	Development Incidence Utility Prevalence Size
X2: Screening	"Screening method" "Screening strategy" "Screening interval"	PMR "Health monitoring" Disease prevention	"Preventive screening" "Medical checkup" "MRI screening" "Total bodyscan" "Total Body checkup" "wb-MRI" "whole body MRI"
X3: Client	-	Customer	Patient Subject Men Women Elderly
X4: PMR	-	-	-

TABLE 4.2: The differenct SPICE dimensions

1. Filter on languages used: German, English and Dutch
2. Filter on year 2000 and later
3. Filter on text availability

Use of search operators

To improve the literature study, search strings can be sharpened and broadened to get more relevant results. In first instance, search terms obliged to be included in the title or abstract form the base of the string. Of all levels within the framework of Table 4.2 terms were chosen, where Boolean operators will be added amongst them. Between two levels often operator "AND" is added, to ensure both terms have been used in the solution space the engine is searching. To broaden or narrow down possible output, terms within other columns can be added or removed, often done by using Boolean operator "OR". By the use of this operator one of all terms included by this operator must be stated in the solution space. At last, terms can end with an asterisk (*), when a word can have several inflections. When this word is cut off by an asterisk, all possible outcomes are still considered.

Different search strings have been tested based on the number of results and scanning the first page to determine the output's relevance.

4.3 Literature search output

Several combination were used, where within the first group of two strings were used to come to the final articles to read. Different definitions of the selected abnormalities were tested based on the number of results and scanning the first page to determine the output's relevance. For the pancreatic and renal body region, same strategy was applied.

Of all three categories of the literature results, studies mainly existed of randomized trials and screening programs in which (often high-risk) study populations were included. Main results and insights used in each paragraph are summarized in the corresponding tables.

In Appendices D.1, D.2 and D.3 the final search strings are given with the flow of article selection per category.

In this section results from the conducted literature search are summarized. After each section of a finding, most relevant information is merged in a global overview of article results, as can be seen in Tables 4.3, 4.4 and 4.5.

Abdominal aortic aneurysm or dilatation

An aneurysm is defined as a local dilatation of the artery caused by a weak spot in the vascular wall of a blood vessel, mostly the aorta. Such a dilatation, or ectasia, is called an aneurysm when this vessel concerns an increase in diameter of 50% compared to the expected value [23][24]. In general, it can be said that the threshold of an abdominal aortic aneurysm (AAA) lies at a diameter of approximately 3.0 centimeter (cm), due to being 1.5 times the normal interval segment of 2.0 cm

According to a research of RIVM, a Dutch institute for public health and environment, it appears that an AAA is diagnosed in 5 up to 10% of investigated men between 65 and 79 years old. Other researches approximate these values, whereas prevalence rates of 4% and 8% are reported. Remarkably, some claim the prevalence to be at a rate between 1-2% in female populations. Besides, prevalence of AAA in men between 50-79 is significantly lower compared to elderly men. These studies all indicate that age and gender are considered as risk factors for the development of an AAA. Furthermore, other relevant risk factors for developing an AAA are -having a history with- smoking hypertension or an ischemic heart disease.

Although patients carrying an AAA are often asymptomatic [23], the consequences cause an increase in importance of guiding patients carefully. Aside from thrombi formation in the lumen or compression of adjacent organs [25], rupture of the AAA is the highest risk to take into account. When a rupture occurs, a mortality rate of 50-80% currently is the standard [25][26].

Focusing on the follow-up and strategy of screening clients with AAA, or an increased risk to develop one, risk factors to develop the abnormality are generally known. Hence, literature suggests the benefits of screening high-risk population yearly to prevent unidentified AAAs from rupturing [27]. According to Longo et al.

patients with a diagnosed AAA are recommended to be followed-up once a year if the AAA has a diameter of >4.5 cm, in 2 years if >4 cm or 3 years when the AAA is >3 cm [25]. The Dutch health institute RIVM recommends follow-up more frequent, namely once a year when between 3 and 4.5 cm, or one every half a year when bigger than 4.5 cm. They advise surgery above 5.5cm when the patient is in good condition [23].

In conclusion, the reason to consider AAA as one of the incidental findings, is the risk of rupture with the related and significantly high mortality rate. Economic evaluations, often performed by a simulation based on scientifically stated probabilities, show that the potential health benefit from screening high-risk populations often outweigh the costs concerned.

Articles included (authors):	***Year of publication:***	***Included information:***	***Details:***
Van Gils et al. [23], Svensjö et al. [27], Sisó-Almirall et al. [25], Meecham et al. [28]	2008, 2014, 2017, 2016	Definition of AAA	
Van Gils et al. [23]	2008	Prevalence	5 to 10%
Longo et al. [24]	2005	Prevalence	up to 8%
Scott et al. [29]	2002	Prevalence	6%
Halett et al. [30]	2000	Prevalence	4%
Sisó-Almirall et al. [25], Benson et al. [31], DeRubertis et al. [32], Jawien et al. [33]	2017, 2016, 2007, 2014	Risk factors	Age
DeRubertis et al. [32], Sisó-Almirall et al. [25], Scott et al. [29]	2007, 2017, 2002	Risk factors	Lower female prevalence
Benson et al. [31], Jawien et al. [33], DeRubertis et al. [32]	2016, 2014, 2007	Risk factors	Smoking
Benson et al. [31]	2016	Risk factors	Hypertension
Jawien et al. [33]	2014	Risk factors	Ischemic heart disease
Svensjö et al. [27], Olchanski et al. [34],	2014, 2014	Risk factors	Cost-effective in high risk population

TABLE 4.3: Overview of most relevant information within literature results (aortic findings)

Pancreas

Because pancreatic abnormalities do not develop symptomatic, most cases are developed to an advanced stage when diagnosed [35]. Therefore, surgery is often no longer an option which causes the high mortality rate and low survival of patients diagnosed [36]. When it considers preventive screening of pancreatic lesions,

specific characteristics of possible abnormalities and their behavior come forward immediately. Abdominal screening of asymptomatic people to detect these lesions, is a way of preventing early-phased tumors for example from developing any further. Additionally, resection of tumors in an early phase has the greatest potential for longer survival [37].

Structure of pancreatic cystic lesions, combined with their size, are the strongest predictor for malignancy and are thus related to survival of the patient [38]. If this size exceeds 2 cm, characterization of the lesion is preferred to the advice of a single follow-up [39].

Since abdominal studies through MRI confirm an incidental prevalence between 13.5 and 19.9% and an increasing incidence, due to an aging population and improvements of imaging, importance of surveillance management rises [39][40]. Screening asymptomatic clients requires knowledge and clear thresholds to determine optimal follow-up.

Hereditary conditions and familial clustering are often listed as main risk factors to develop pancreatic tumors. Additionally, clinical predispositions due to alcohol, age, smoking history and diabetes are proven as risk factors as well. Due to a relatively low prevalence but high potential consequences, screening high-risk population is often recommended [37], [40]–[42].

Articles included (authors):	***Year of publication:***	***Included information:***	***Details:***
Lee et al. [35]	2009	Prognosis	Relevance of early diagnosis
Chang et al. [36]	2014	Prognosis	Relevance of early diagnosis
Bruenderman et al. [37]	2015	Prognosis	Relevance of early diagnosis
Chin Hur et al. [38]	2017	Malignancy predictor	Structure and size of the lesion
Gore et al. [39]	2012	Prevalence rate	13.5-19.9%
Bruenderman et al. [37], Becker et al. [40], Pezzilli et al. [41], Poruk et al. [42]	2015	Risk factors	Alcohol, age, smoking, diabetes

TABLE 4.4: Overview of most relevant information within literature results (pancreatic findings)

Kidney

Compared to the two other abnormalities within the scope of this report, prevalence of abnormalities found in kidneys are relatively high. Additionally, in contrast to AAAs being a specific finding, the kidney can include several abnormalities. These incidentalomas, often referred to as renal masses, can be identified through MRI.

Concerning MRI capabilities, kidney failure for example has not been considered at all, since this disease is diagnosed based on the glomerular filtration rate (GFR) and blood pressure [43].

This high prevalence of identifying renal masses is 10.7% (cystic lesions), based on asymptomatic populations from [44]. Additional research showed that 12.3% incidentalomas were diagnosed within a research population of 6678 asymptotic elderly people. Of these findings, 9.4% were labeled as cystic, and 15 (0.22%) renal cell carcinomas (RCCs) were identified [45].

Since the size of a renal cyst for example is not considered as most relevant indicator for malignancy, increasing complexity of the cyst is used as preferred classifier within the Bosniak classification system [46]. Besides, survival rates of patients with normal cystic lesions for example, are relatively low. Severity increases when these lesions are behaving malignant, and chances are there to influence patient's health states significantly [46].

Incidental tumors detected by preventive screening have a better prognosis and provided longer disease-free survival [47]. Combined with a rising incidence of renal cell carcinoma over the years, this urges the relevance of abdominal screening and surveillance of renal incidentalomas [48].

Furthermore, developing renal masses of which specifically cystic lesions, are positively related to age, gender and smoking. People with renal stones and serum creatinine are considered high-risk as well [44].

Articles included (Authors):	***Year of publication:***	***Included information:***	***Details:***
Malaeb et al. [45]	2004	Prevalence rate	12.30%
Chang et al. [44]	2007	Prevalence rate	10.70%
Chang et al. [44]	2007	Risk factors	Age, gender, smoking
Sohaib et al. [46]	2012	Follow-up	Bosniak classification

TABLE 4.5: Overview of most relevant information within literature results (renal findings)

Summary of gathered literature

Since screening techniques are improving and become more capable in detecting several abnormalities, the number of these outcomes and incidental findings are increasing. Although these advanced techniques nowadays cause a higher incidence or detection of new abnormalities, it ensures early diagnosis and thus the possibility to treat this finding. In all three selected "cases" in this report, early asymptomatic diagnosis often yields a better prognosis for this specific person [49]–[51]. An interesting comparison when thinking of all abnormalities not screened at all. However, a new challenge arises of managing the appropriate follow-up and surveillance for all

different sorts and behaviors within the diagnosed incidentalomas [49]. Percentages from researches in Japan for example, add up to a prevalence of abnormalities found in approximately 40-50% of the population (slightly more in men; 44.5% against 34.2% in females) [52][53]. Consequently, the relevance increases of managing these findings and select or filter all potential malignant ones requiring follow-up.

In line with results from this literature search, some scientific thresholds and guidelines are gathered and analyzed in a report of the American College of Radiology. Some of these flowcharts to support medical decision making for surveillance are added as appendices [54].

Conclusion

In conclusion, this literature study have lead to many relevant insights, focusing on different characteristics of the selected finding within the abdominal body region. As described previously, many studies have been performed to determine risk factors for developing specific diseases or suspected masses. Characteristics of high-risk population have been defined, and aside from proving the effectiveness of screening them, prevalence and some incidence rates were presented as well.

On the contrary, little data is available on best practices and determined screening strategies for these populations, especially when looking at the general "healthy" population. Reason for this can be explained by the 12 criteria for performing preventive screening and checker whether it is worthwile, established by Wilson and Jungner in 1968 [55]. Within these criteria, the combination of health care costs and the effectiveness of performing screening, is highly relevant. Screening all people within a general population will increase health care costs immensely, where on the other hand few detections of abnormalities will occur. Low prevalences of meaningful findings in average risk persons would offset the cost-effectiveness ratio [56]. Therefore, most studies do recommend preventive screening, but mostly focusing on high-risk populations [27][51]. Consequently, risk factors in developing the selected potential outcomes are important to take into account when considering a specific outcome such as an AAA or a pancreatic (or renal) lesion.

Before next Chapter 5 further treats these literature insights compared to the gathered company's data, an extra literature search will be performed, focusing merely on development and incidence of diseases occurring, to get to strategies for determining a more personalized recommendation for screening repeatedly.

4.4 Personalized follow-up

Since valuable literature has been gathered, but this fourth research question could not have been answered so far, an extra search will be performed. To come to a screening recommendation for clients potentially developing an abnormality in the future, a comparison is made from scientific research, of how to achieve this personalized follow-up strategies.

Literature examples

As a result of the previous literature study, focusing solely on the selected abnormalities, few literature was found concerning strategies and techniques to get to a personalized follow-up. With the results of the initial search and the following additional search in this chapter, no specific techniques or models were used to come to a personalized follow-up: Usually, medical research found are performed within a prospective cohort, and using for example the setup of a randomized controlled trial (RCT) several outcomes, risk factors and probabilities for diseases can be concluded.

In most of the articles found in the previous literature study, a randomized controlled trial was the initial study setup. Based on these results, risk factors for developing diseases are determined. But, this is based retrospectively on a preselected population. When advising clients for future follow-up, this is just an indication. As highlighted in the article of Ladd et al., these preventive research setups can be divided into two groups of patients, namely asymptomatic patients or patients with risk factors [16]. But no researches were found in which existing data of an asymptomatic or general population was used in a way this report does.

Therefore, a search is performed focusing on the advised and personalized intervals of it based on different possible criteria. Most important question to be answered by doing so, is what kind of approaches can be used to optimize a personalized revisit recommendation.

Search strategy According to Wohlin et al. [57], it is important to first gather a relevant start set of articles. After doing so, backwards snowballing is applied, in which the reference list was used to search for potentially relevant studies. To come up with information comparable to this report's scenario, relevant terms and definitions from known articles were used to broaden the range of the theoretical framework. By using these terms as input in multiple article databases (Scopus and Pubmed, see 4.6), it resulted into the article set of table 4.7. Second part of this table includes the first iteration of backwards snowballing.

Database used:	***Search term used***	***Results***
Scopus	TITLE-ABS-KEY (risk-based AND screening AND follow-up)	138
Pubmed	(risk-based OR personalized) AND (follow-up OR recommendations OR strategies)	7866
Pubmed	(risk-based OR personalized) AND (follow-up OR recommendations OR strategies) AND interval	315

TABLE 4.6: Search input for personalized screening literature

Articles included (authors):	***Year of publication:***	***Study objective:***
Dong et al. [58]	2017	Risk stratification, cumulative risk
Hostetter et al. [59], Witteveen et al. [60]	2018	Personalized follow-up
Ripping et al. [61]	2016	Personalized screening, cumulative risk
Added through backwards snowballing:		
Tammemagi et al. [62]	2011	Risk prediction
Ayer et al. [63]	2012	Personalized screening
Zhang et al. [64]	2012	Screening policy optimization
Otten et al. [65]	2018	Stratified follow-up

TABLE 4.7: Search results for personalized screening literature

Results

To reflect data insights from a population collectively, towards an individual optimized interval, a Markov model used as in Witteveen et al. [60] is a way to approach this objective. As described in this research publication, large datasets were used of comparable patients of which all follow-up moments were registered carefully. With this registration, multiple risk factors were included, to see through logistic regression models to which extent they influence the decision of shortening the interval of treatment or not.

On the contrary, the report's current dataset does include a huge population, but when focusing solely on the clients with an abnormality diagnosed, only 414 situations remain. Of these, related age and gender are known. Looking at the development of a disease, and thus information of clients over multiple moments of screening, an amount of 118 were included. Of this group, 49 cases had a screening without a diagnosis prior to developing this abnormality officially. To proof a statistically significant optimization for individual clients based on this strategy, more data would be needed.

Furthermore, as applied in these models, specific criteria are used to come to an optimal solution. Within Markov models used for health care sector's purposes, this often is related to a trade-off between quality of life gained and costs concerned. As in the situation of Prescan, and the earlier stated requirements of Wilson and Jungner [55], this cost-effectiveness decision or threshold, is different when the client is paying him or herself. In the end, they are the ones determining whether or not there is willingness to pay for preventive screening. But, as seen in the model of Otten et al. [65] another criterion is used to optimize, namely minimizing the probability of a recurrent tumor. A trade-off is made between life years gained, and the disutility of performing an MRI scan when the decision is made to apply for screening, defined

as a reduction in the reward gained. This way of decision making will not make significant differences in the outcome when applied to our own scenario, due to the lower disutility of preventive checkups, when MRI techniques are used. Physically performing an MRI is not harmful, aside from potential psychosocial consequences of applying screening.

Therefore, this report initially focuses on estimating the risk of developing clinically relevant findings, to the extent in which that is possible with the available data.

4.5 Conclusion

Initially, a literature search was conducted focusing on indicators for disease development and screening trends from the perspective of the selected abdominal MRI findings. Since this gave relevant insights, but was not sufficient in information provision to take further steps in the estimation of risks in developing clinically relevant findings in future, a second search was performed, with a better focus on solely strategies for follow-up determination. By applying these methods extra insights in both directions were discussed. What will be discussed hereafter is explained in next paragraphs.

First, Chapter 5 discusses, based on a comparison with measures from literature, the findings detected within the retrospective Prescan cohort. With this comparison, characteristics of recurrent client's findings can be categorized, assuming them to be earlier staged then abnormalities found within the regular health care.

Then, there will be worked towards a generally stated interval strategy in undergoing preventive MRI, by relating and comparing different screening intervals to the outcomes identified. Since the answering of research question four has been reconsidered, the potential of optimizing a personalized recommendation for screening is identified. Although current data does not enable determining an optimized follow-up yet, with the strategy used in upcoming Chapter 6 a model will be setup out of subgroups within the gathered company's dataset. With this model there will be worked towards calculations of estimating risks of developing clinically relevant findings.

Chapter 5

Analyses of the data selections

5.1 Categorizing Prescan's findings

To work towards the determination of screening strategies per client, the main objective is to get to an advice for a personalized interval of performing MRI research. There exists a lack of literature concerning these strategies, but several characteristics per finding have been found. These will be compared to outcomes of the data preparation and analysis in Chapter 3, and literature results of Chapter 4.

Main characteristics and thresholds

Findings of MRIs performed at Prescan are analyzed and registered in a report by the involved radiologist. To compare and validate the outcomes of the gathered data, in this report these findings will be categorized. This is based on the most relevant characteristics per finding, following from the conducted literature search. In addition, categories used within the Aftercare department and the earlier conducted research in Chapter 2 are partly used as well.

Within the potential AAA group of findings, age, sex and size of the detected dilatation were labeled as important risk factors in developing an AAA. According to literature and the ACR, development in size is a relevant risk factor to track over time as well.

For the second group, this includes pancreatic lesions like a cystic mass. Of these findings, it's important the following risk factors are identified: size, contour and location, and whether the mass can evolve to a more malignant stage like BD-IPMN, SCN, MCN or PDAC.

When renal masses are considered, its categorization is mostly based on the behavior of the specific abnormality, where a hemorrhagic cyst or calcification of it are good indications for their stage in the Bosniak classification system. Furthermore, size is partly taken into account, where cysts from 30mm are included for follow-up.

Overview of clustered outcomes

All three subgroups considered within the scope of this research include clients for whom an abnormality has been diagnosed and who are recommended to perform

follow-up research. Simultaneously, as stated in Chapter 3, the definition of a finding equals a screening outcomes of which follow-up investigation is advised to reduce the risk of getting complications.

As a result, following from the data gathering and preparation methods previously, it appeared that 414 of all clients undergoing MRI in the stated period were diagnosed with one of the included findings. Within all three groups of the selected body regions, an outcome is labeled as finding following the definition from this paragraph. From this perspective, outcomes included are all defined as clinically relevant. When for example a renal or pancreatic cyst is included in the data set, it means that the composition or behavior of this finding deviated from what was expected. Some potentially relevant findings must be followed-up in future, so in case of further development such as progression or dangerous this can be identified.

Unavailable section *This section is not available in the public version.*

Unavailable section *This section is not available in the public version.*

Unavailable section *This section is not available in the public version.*

5.2 Conclusion

This chapter shows a clustering of the company's data gathered, based on scientific definitions and thresholds from literature. By doing so, it is shown that of the relevant findings included, characteristcs like size, location and behavior of them can be categorized in line with these definitions. Although the discussion of managing and screening for clinically relevant findings remains, and depends on its severity and potential harm, the following can be concluded: findings included in the substantiated selection of this data setup, on average are earlier staged compared to findings diagnosed symptomatically in regular care.

Within upcoming Chapter 6, a model for estimating risks of developing findings is conducted, based on the second part of the literature search insights. By doing so, all findings found are included, for both the revisiting clients and those who visited Prescan once during the research period determined.

Chapter 6

Model description estimating risks

After the data anlysis performed in Chapter 3, all the client's ages and their development through multiple researches have been visualized, in which several selections of findings were highlighted. Characteristics of the selections with all findings were presented separately. Afterwards, all findings were generalized into the definition of a clinically relevant finding, for which follow-up has been advised by the treating care providers. With these visualizations, an overview is given of what the ages per related research moments were, and what the personalized care path has been per client at Prescan.

To come to an estimation for new clients' potential revisit recommendations, a model will be produced to estimate the risk of developing a clinically relevant finding. This will be done by the use of the available and prepared data, for several selections of client groups. Furthermore, the fact that this report considers research in the field of health care technology and management, must be taken into account as well. Therefore, some assumptions are made tending towards this field of knowledge, which can slightly influence the clinical appropriateness in some situations.

Together with the previous literature and outcoming results of the modeling subgroup calculations, fourth research question is treated: *Is it feasible to accurately determine the probability of developing and growing a clinically relevant finding based on currently available literature and data, and if so, can this be translated towards a personalized interval?*

6.1 Risk estimation

According to MacMahon and Pugh [66], epidemiology is the study of the distribution and determinants of disease frequency in human populations. Epidemiology as we see today, is defined as the study of the distribution of health-related states and events in populations. The objective of many of these studies is to obtain a valid and precise estimate of the effect on occurrence of disease. In the setup of the model, the outcome can be seen as binary, where no diagnosis or recommended follow-up is seen as the group of clients *without a finding*. After diagnosing an abnormality, and thus when follow-up is required, this counts for an event within the *finding* group.

In the working field of healthcare research, the medical term prevalence refers to the proportion of persons in a population who have a particular disease or attribute at a specified point of time, or during a period. It gives information of the amount of people having a certain disease, and for example how widespread it is.

To estimate the risk for developing a clinically relevant finding of a returning client at Prescan, data from comparative clients will be analyzed to come to a recommendation for follow-up. Since this data already exists over a large population, initially the research moments of all included clients are sequenced. Of each of the individuals it is known per corresponding scan moment, whether they had been diagnosed with a finding or not.

Considering the objective of these prospective risk estimations, it is even more interesting and required to look at the risk of developing a disease in future. According to Witteveen and Otten [60][65], probabilities of state transition from the initial state (healthy) to the next (diagnosed finding) have to be known. But, in this situation, asymptomatic clients aren't screened every now and then, and certainly not all of them performed a screening multiple times. This occurrence of having unknown parts of information within the research population, is often called censoring. To account for this, health scientists got to use a more adequate way of describing this estimation of developing a disease; the incidence.

Incidence rates

When looking at a period of ten years, for example between the age of 60 and 70, some clients possibly visited Prescan multiple times. Assume these visits took place at ages 61, 64 and 67, and no finding was diagnosed. For this specific client, 6 information years are known within this period of ten years. Another person, diagnosed with a disease at 65, but declared healthy at 63, solely has these two research moments as measures at Prescan. Both clients are censored subjects, because they are only partially observable during the stated period of ten years. Using the definition of the prevalence, one could say that the risk of developing the disease is 50%. But when measuring the frequency of disease occurrence in a population, it is insufficient to merely determine the prevalence. When looking at the incidence, one takes into account the extent to which new cases of certain diseases will develop. Therefore the amount of person years with information about having the disease or not, are taken into account. Relatively speaking, this gives a better risk estimation of developing the disease.

According to Rothman et al. [67], the measured incidence can be presented through the incidence rate (IR). This rate is defined as a quotient, of which the numerator includes all events occurring during the specific period of time, and the denominator summing up over all information years every participating individual

is included (person years):

$$Incidence\ rate = \frac{Number\ of\ disease\ onsets}{\sum Person\ time\ spent\ in\ population} \tag{6.1}$$

The incidence rate conveys information about the risk of contracting the disease, whereas prevalence gives an indication of how widespread it is. It determines a probability of developing a specific disease over a stated period of time. Therefore, it can be seen as an estimate of risk.

6.2 Modeling IR development per subgroup

In this section the setup of the model will be highlighted. This model represents an automation of calculating incidence rates of preferred subgroups selected. Of the whole research population clients over the whole period determined, based on their age, subselections can be created. Based on filtering out the selected range of clients' ages at moments of their screening, groups are created and sequenced, after which the related incidence rates are presented.

After gathering all model outcomes, cumulative incidences are used to determine risk estimations when years of a potential revisit are preferred to determine.

Dataset description

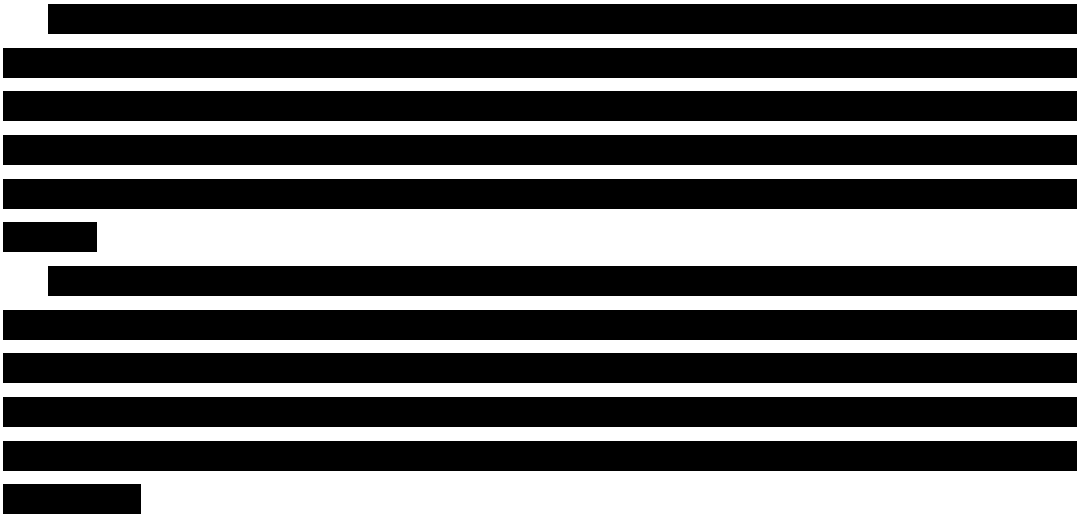
In order to get to an understanding of the model, first the input data will be described.

The data selections and visualizations in Chapter 3, give relevant insights about the general Prescan population and add detailed information between the group of clients with a finding. But still, the question arises of what recommendations can be given to clients for a future return for screening. This is, of course, with the objective of outweighing pros and cons, and estimating the risk of developing such a clinically relevant finding in the meantime. Clients have to consider such an estimation with possible consequences of a screening (costs, emotional effects), and probabilities of not detecting potential findings when not apply for screening. The first part of these analyses was focusing on different possible outcomes of the screening sessions. For each outcome one can imagine the follow-up and especially future consequences differ per client and scenario. Since clients coming to Prescan are asymptomatic and do not know what finding they will face, from now on all three findings categories will be generally taken into account as a whole. In this way, the added value of recommendation as prediction for an interval strategy will be increased.

Furthermore, all data of the same research population without having one of these findings, are added as a whole for the input dataset. They form the counterpart of clients with a diagnosed finding. By doing so, a next step is taken towards the risk estimation of developing a relevant finding, because a comparison can be made among these subgroups and their characteristics, of who did or did not have

a finding. With this estimation, an extra support for clients is created, to help them determine a scientifically appropriate interval to return for preventive screening.

By selecting all clients undergoing a TBS in between October 2009 and May 2018, the whole dataset is labeled. From this, all earlier described clients with a recommendation to perform follow-up research based on their findings, are excluded to ensure this is the group of clients who did not have one of the selected findings.



Model setup

To come up with the number of cases (events) as ratio of all person years, initially the period in which will be searched is determined. When a client is for example 62 at moment of screening, this person is included when the initial period of LB and UB equals [60-70]. Since the LB and UB determine which clients are included in this calculation, it means that every client within this selection visited Prescan at least once in between this period of [LB,UB]. However, this does not mean their whole period of their screening path is in between these bounds. Therefore, their share is calculated based on reframing their path and deciding to which extent this part is taken into account.

To actually calculate this, the second step is using the distinction made between the characteristics of the selected clients: whether they were diagnosed with a clinically relevant finding during their screening visits at Prescan, determines which procedures to call. In the preparation of the dataset this characteristic was included, so a simple check can confirm whether a client belongs to the "non-finding" or "finding" group.

Group of clients without an event For the clients in this group, methods differ based on whether they visited the screening clinics one single time, or multiple. The distance (period) between the last and first visit is indicator for this repeating character; a difference bigger than zero means there are more research moments than just a single one. Secondly, it is important to know what share must be taken into account for this group of clients. Since the objective is to determine the number of

years in which it is known that this client was not diagnosed with a specific finding, the moment of their last screening is an indicator for this calculation. When it is determined no finding has been diagnosed, the period between this moment of screening and the stated LB is calculated. In case of a known last scan which took place after the stated UB, the whole distance between the current UB and LB counts as share of the clients with no finding.

Group of clients with an event In this second group of subjects with a diagnosed finding in at least one of their screening moments, same procedures as in the first group are used. Furthermore, within the definition and use towards the incidence rate, one important characteristic is taken into account when calculations take place; after the moment of screening at diagnosing the specific finding, further person years of this subject are not taken into account anymore. In general this is because further processes of clients with a finding differs from others, due to potential treatment procedures or shorter periods of follow-up for monitoring disease development.

Therefore, a distinction between these groups is made initially, so that a summation can be performed over both groups, while taking care of this difference in accounting.

Calculating the incidence According to equation 6.1, the numerator of the formula represents a count over both the earlier described groups of research subjects. Within these groups, a check is performed on whether a finding was diagnosed or not. Subsequently, the denominator keeps track of the total amount of person years that each of the individual clients were taken into account. All possible situations of revisiting clients will be explained in the next paragraph, where the technical properties of each calculation are highlighted.

Technical setup

In this paragraph each step of the model towards IR calculation is explained. Starting with the initialization of the model, followed by an explanation per subject group and results presentation.

Initialization Before each of the selections performed, the developed initialization procedure is used to clear all earlier results and set the new lower upper bound for the next selection. In the overview of the dashboard including all steps, this so-called reset button can be found at the left upper corner. Furthermore, while initializing each step, lower and upper bound will be determined, and can be decreased or increased within step 1a and 1b. In this way, the preferred bin combination of [LB,UB] is set for the following selection of subjects out of the prepared dataset.

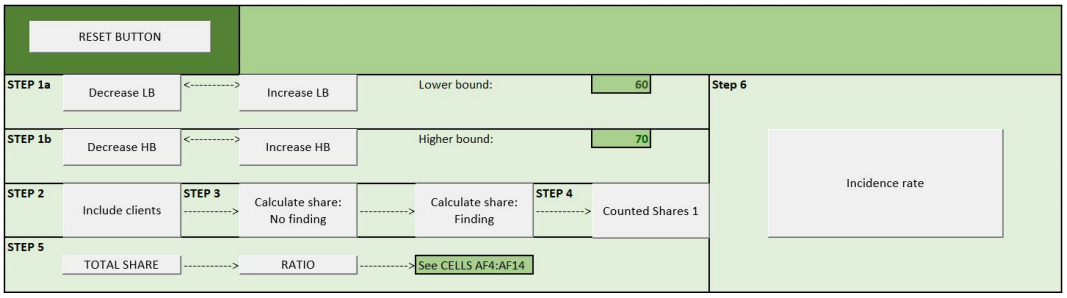


FIGURE 6.1: Model of the simulator and steps taken to calculate the incidence rates

Calculating shares of person years After initialization of the model, some steps are taken to perform the shares calculation of each individual subject. This means, that per subject distances from the relevant moments to LB or UB are used to determine their share in the total amount of person years counted.

First, step 2 represents the main part of each selection procedure, namely selecting all subjects to take into account. With this procedure, all subjects' information strings are included in the newly created selection, if one of their research moments lies within the initial bin of LB and UB. When setting a lower (LB) and upper bound (UB) for this selection, with the use of the corresponding method a loop will be performed, in which all clients of this period are taken into account.

After determining the subset, of all corresponding subjects their share of person years within the bin is calculated. Logically, if all of the screening moments are before or after of this bin, they are not included. For the non-finding group, periods to take into account are as follows:

- If there is a single screening moment (SM), which thus lies in between [LB,UB]:

$$Accounted\ share = \ Age\ at\ SM\ -\ LB \tag{6.2}$$

- If there are multiple screening moments, so before or in between [LB,UB]:

$$Accounted\ share = \ Age\ at\ latest\ SM\ -\ LB \tag{6.3}$$

- If there are multiple screening moments, of which at least one after UB:

$$Accounted\ share = \ UB\ -\ LB \tag{6.4}$$

When the model loops over all subjects with a finding, their share is calculated like this:

- If moment of finding lies before [LB,UB]:

$$Accounted\ share = \ 0 \tag{6.5}$$

- If moment of finding lies in between [LB,UB]:

$$\textit{Accounted share} = \textit{Age at moment of finding - LB} \tag{6.6}$$

- If moment of finding lies after [LB,UB]:

$$\textit{Accounted share} = \textit{Last performed SM - LB} \tag{6.7}$$

Incidence rates After passing through step 3 of the model, in which for both groups all relevant distances of time periods are highlighted, steps 4 and 5 are conducted. In this part, all individual shares of person years are determined scenario-based. This will be explained in the next paragraph in more detail. Then, all shares are summed to get to the collective total of both groups, equal to the denominator of equation 6.1, the incidence rate.

To achieve an IR for the selected population, of all included subjects an extra method counts the number of events taking place within the period of the selected [LB,UB]. Counting this number of events occurring within the period, will lead to a total being equal to the numerator of the IR.

Since the IR can take a myriad of possible numbers, epidemiologists use it in a way that the severity of the risk is tangible immediately. Therefore, in this report we use a rate in which the number of events per thousand person years is presented, by multiplying the number of disease onsets by thousand. When for example 2 events occur over 400 person years, the IR will have a value of 5 events per 1000 person years (because 2 times 1000, divided by 400, equals 5).

6.3 Scenario analysis

Following the exact definition of calculating the incidence rate of a selected population, we assume the distribution of events occurring is dichotomous. This means that an event occurs or not, where the model accounts a value 1 or 0 respectively. Of a client undergoing a TBS having no diagnosis of a clinically relevant finding at all, from this moment it is known that no finding was found. From this moment as well, further development is not certain to predict. Therefore, advice for a potential revisit is needed based on comparable clients of which we do have information. However, the model in first instance assumes a finding to develop from 0 to 1 in a split second of time. In real human bodies, developing such a finding from an initially healthy state, occurs over time. Although we do not know exactly when this event occurs, due to not tracking every subject during the intermediate period.

To account for this uncertainty, a scenario analysis is performed, in which different assumptions of developing a finding are compared.

Scenario 1: Immediate development As introduced in this section, the first scenario is the base situation when using the definition of an incidence rate. When calculating the share of all included subjects, highlighted formulas from the previous section are used (6.2 up to 6.7).

If there is only one moment of screening within the stated period, passed time from the beginning of this period until that screening moment is calculated. In case of multiple researches, the age of LB will be subtracted from the latest screening moment.

When it is known from data that some subjects will develop a clinically relevant finding in future, but the related moment of screening lies after the specified UB, this scenario does not count any of its development as an event. This forms the base solution, where events are solely calculated as 0 or 1, when taking place within period [LB,UB]. Therefore, latest negative screening moment (in which no diagnosis was made) is labeled as value to subtract LB from to calculate the included person years.

Main assumptions in this scenario are thus that the development of the disease is not taken into account, besides the fact that an event occurred when the screening took place. In this case, the health state of a client changes from "healthy" (0) to "unhealthy" (1).

Scenario 2: Linear development With the perspective of this scenario, we assume that clinically relevant findings develop linearly and relatively slow over time. For the setup of the model, formulas resulting in the person year summation for the non-finding group will remain the same as scenario 1. When a finding is diagnosed somewhere during [LB,UB], it is assumed this finding developed linearly over time. Since the moment of diagnosis lies in between the current bin, total sum of events taken into account always equal one. But, the main difference in accounting person years compared to scenario 1, is that subjects of which we know an event will occur in future, are fully included. Therefore, equation 6.7 is adjusted to 6.8, for calculating the amount of person years included:

- If moment of finding lies after [LB,UB]:

$$\textit{Accounted share} = \textit{UB - LB} \tag{6.8}$$

When considering the count of number of events occurring in this scenario, calculations change as well. Where in scenario 1 events are solely counted as a whole when diagnosis takes place within [LB,UB], the difference in this case is the assumption that a clinically relevant develops linearly (relatively slow) over time. Therefore, for clients within the current bin not being diagnosed with a finding yet, the probability that they developed a finding after a negative screening (no finding) but within the stated period is added to the sum of events partly. Which part must be taken into account, depends on the time period between the last negative screening moment,

UB, and the MRI screening in future where the diagnosis of an abnormality is made. To calculate this, a method based on the triangular distribution is used, according to Kotz and van Dorp [68].

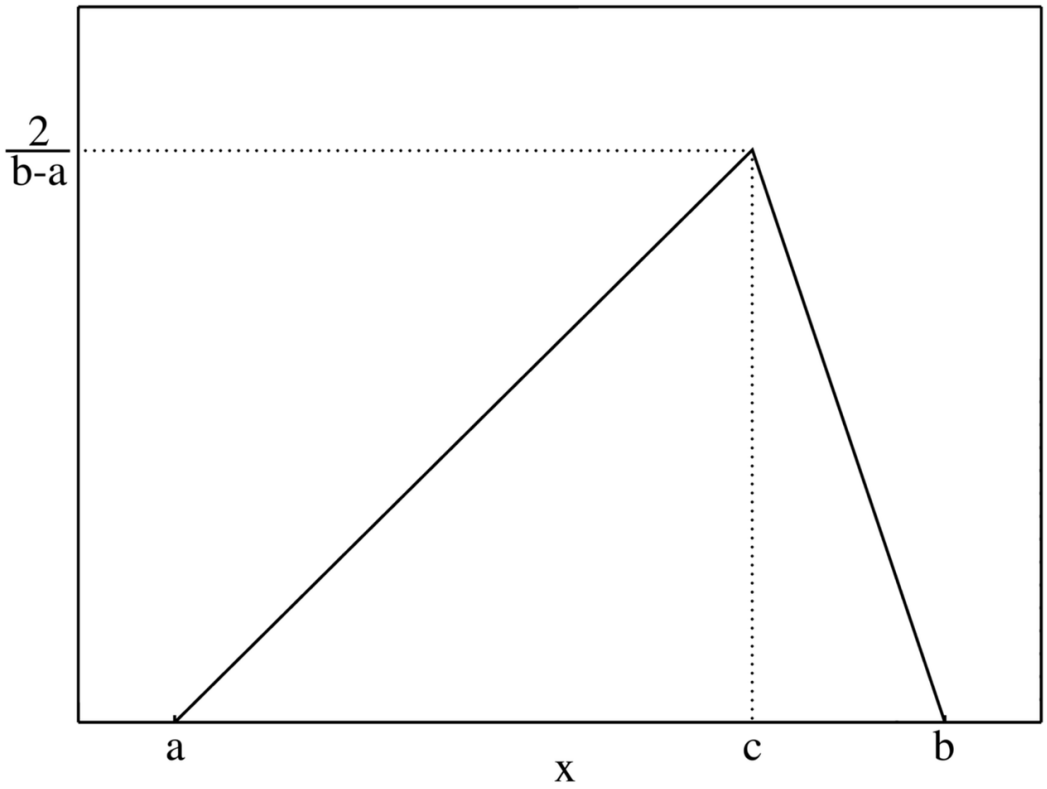


FIGURE 6.2: Probability density function of a triangular distribution

In Figure 6.2 the probability density function (PDF) of a triangular distribution is given. Reflecting this to this report's situation, values a and c equal the start of the unknown period after the latest negative screening before UB, and the point in time after UB when diagnosis takes place. Assuming this finding develops linear over time, the probability it occurs before [LB,UB] after the negative screening, is a cumulative probability that develops linear from 0% to 100% (which equals 1). In this case, x equals UB for the current selection.

The following graph of Figure 6.3 belongs to a cumulative distribution function (CDF) like this. Since a probability to develop an abnormality increases from 0 to 1, this scenario ends when x approaches and eventually equals c.

Within the theory of the triangular cumulative distribution, the following formula is used when $a < x <= c$:

$$CDF = \frac{(x-a)^2}{(b-a)(c-a)} \tag{6.9}$$

Since x never outweighs c, because mode c equals upper limit b in this scenario, it can be said that *(b-a)* equals *(c-a)*. Translating this to our simulation model with UB

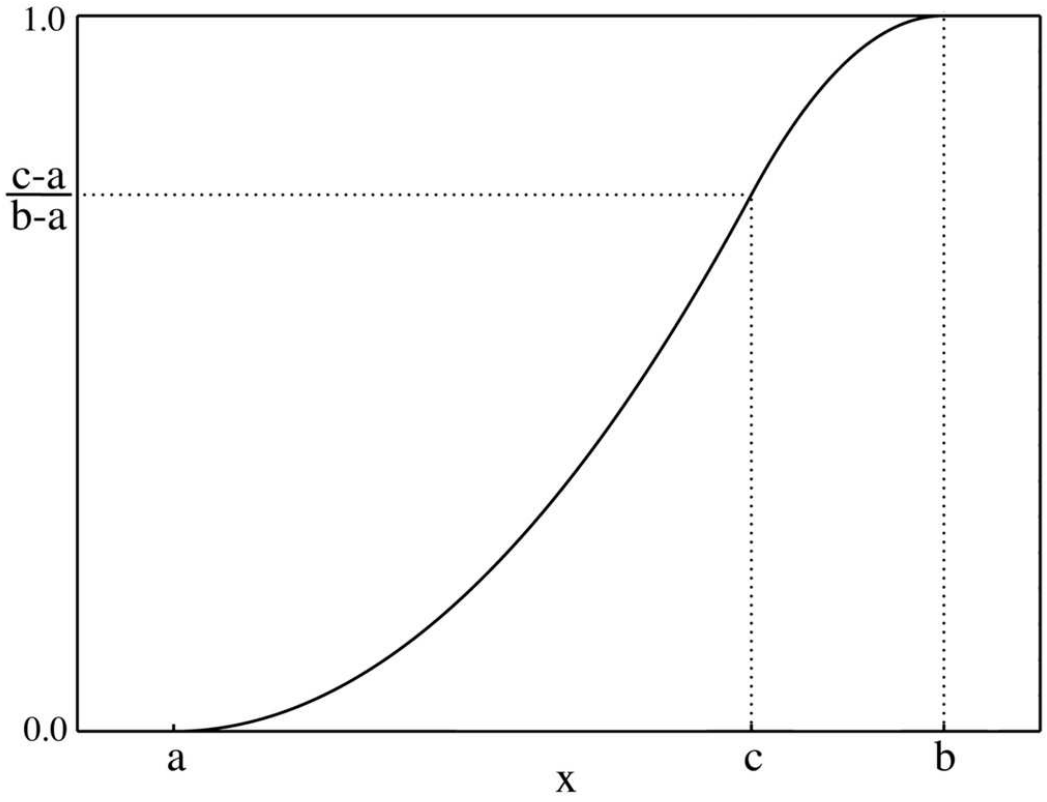


FIGURE 6.3: Cumulative density function

and the screening moments with and without a finding ($SM_{Finding}$ and $SM_{NoFinding}$) results in the following:

$$CDF = \frac{(UB - SM_{NoFinding})^2}{(SM_{Finding} - SM_{NoFinding})^2} \tag{6.10}$$

Scenario 3: Quick linear development This third scenario assumes that a finding develops linear but quickly over time. From this perspective, the probability that within the selected period the future event already occurred follows a cumulative probability from 0% to 100% as well. The only difference in the assumption of a quick development, is that findings usually do not develop that long before a diagnosis, otherwise it would have been notices. When far ahead of the moment of such a diagnosis, the finding could not have been diagnosed at all (due to non-existence or early staged development). Assuming this, scenario 3 therefore considers this linear development only applies if the screening moment at one year before diagnosis ($SM_{Finding}$ - 1), lies before UB.

To apply this to the model if such a situation occurs, equation 6.10 is adjusted as followed:

$$CDF = \frac{(UB - (SM_{Finding} - 1))^2}{(SM_{Finding} - SM_{NoFinding})^2} \tag{6.11}$$

The period for quick development of a finding is debatable and depends on multiple criteria (individual risk factors, severity of the disease etc.), in which it can easily be changed to two years for example. If so, the numerator of equation 6.11 will increase due to a longer period of potential development of a specific finding.

6.4 Model verification and validation

Since we have been able to setup the model as described, verification and validation of the model should be involved, since this is a significant element of any study in which a model is setup. Without thorough verification and validation there are no grounds on which to place confidence in a study's results.

Verification

According to Robinson et al., verification is the process of ensuring that the model design has been transformed into a computer model with sufficient accuracy [69]. Since this model merely represents a sequence of calculations of subgroup selections out of the research population, the only thing to verify is whether steps are computed carefully. Therefore, the model is created in a way that several steps are computed separately, and eventually are combined to calculate the objective of this model; the incidence rates. Each time an [LB,UB] combination is selected, the subgroup considered and its corresponding information is stored at a different place. Each next step is stored separately as well, when for example person years per scenario are calculated. By doing so, different stages towards the model outcomes were checked manually, to ensure the steps were implemented with sufficient accuracy.

Validation

Looking at the validation of a model, it is related to verification in a way that verification can be seen as subset of the wider issue of validation. A key concept of this broader validation is the extent to which it serves the specific purpose, or objective of the setup. With the use of these definitions, this report makes a distinction between concept and data validation, and validation of the solution.

Concept and data validation First, the concept validation in this report's situation is different compared to standard validations of simulation models for example. This is, because this model's objective is solely to dynamically determine incidence rates over different subgroups selected. To show sensitivity of the outcomes and influence of using the definition of IR, multiple scenarios are considered.

In conclusion, this report's model can be considered valid, due to the fact that the model serves the objective to present incidence rates aligned with its definition and calculation. On the other hand, it also is completely depending on its data input. For example, next chapter shows results with [LB,UB] differences of two years, because

a smaller bin size would be too dependent of events occurring by chance. With an average disease prevalence of 1% for example, outputs can easily differ too much, when a small amount of clients are included in the subgroup (n=50 for example).

Solution validation Validating the solutions of the model based on outcomes presented in upcoming Chapter 7, is achieved by comparing these outcomes to real system measurements and experts' intuition. For example, when looking at the incidence rates of diseases calculated between zero and 25 per 1000 person years as first results will show, this means over a period of one year zero to 25 out of 1000 subjects will be diagnosed with this specific disease. These solutions can be considered as reasonable values.

Strengthening these statements, can be accomplished by feeding the subgroups and related calculations with more and different data. For example, when we want to improve the confidence of the model to proof validation, data of other findings and body regions can be gathered and used as input of the same model. By comparing differences in incidence rates, and fitting these outcomes to existing data on disease incidence, external validity of the model can be shown.

6.5 Conclusion

This chapter explained how and what type of calculations were used to come to the model presented. With the use of several scenario setups, differences in perspectives of the incidence rate definition are compared. Studies focusing on incidence of diseases, often are prospectively conducted, in which fixed periods of research are used. In our model, combinations of [LB,UB] ages can be considered as multiple cohorts of which information is gathered. Since some screening moments of clients are known, but are not included due to these age boundaries, the scenarios are implemented to show potential influence. Validation of the model outcomes is not yet as preferred, with the amounts and type of data currently available, but can be valuable when future research is applied.

Next Chapter 7 will present the outcomes of the calculations, and gives insights in how to recommend potential revisits based on currently used data.

Chapter 7

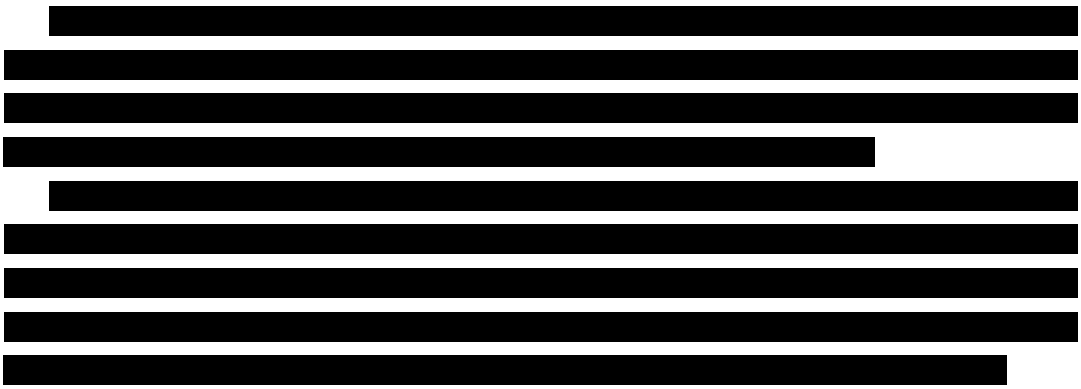
Results

This chapter elaborates with the previous Chapter 6 in which a model was created, by which several selections of subjects were assessed. All calculations work towards an overall risk estimation of developing abdominal findings , based on retrospective data collected from Prescan's MRI scans.

In the first part of the results chapter, a general analysis will be presented where insights over the whole research population are discussed. Subsequently, more detailed selections and predictions will be highlighted.

7.1 Development of incidence rates

When running the model with an initial [LB,UB] combination of ages 20 to 30 and an increase of 10 years each calculation, the output of the simulation model is given in Figure 7.1. The related incidence rates as output over the ages 20 to 90 are added in Appendix Table F.1. Age group 90-100 was excluded due to the fact that only two subjects were included.



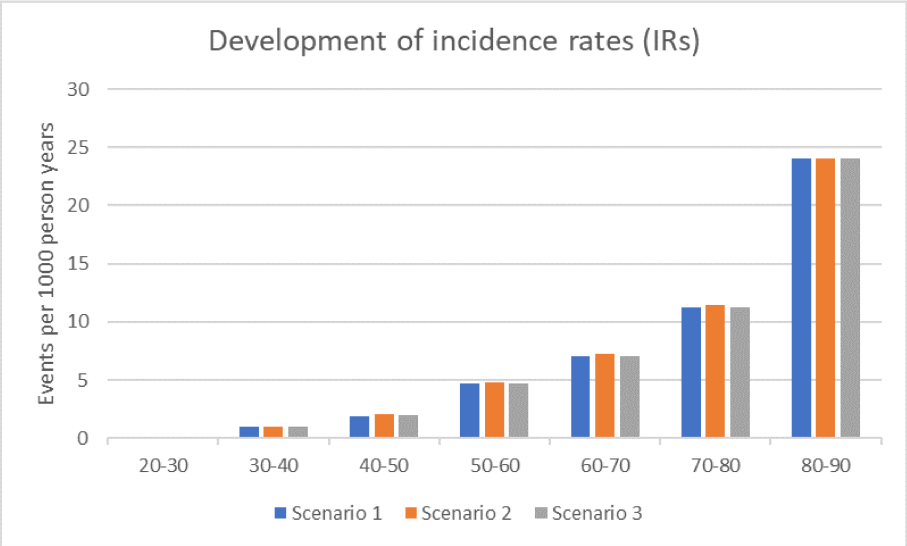


FIGURE 7.1: Incidence rate development between ages 20 and 90 (bins groups of 10 years)

When focusing on the majority of the subjects between 30 and 80, an increase in [LB,UB] of 5 years is used, which can be seen in Figure 7.2 (corresponding numbers of this graphs can be found in the table Appendices F.2).

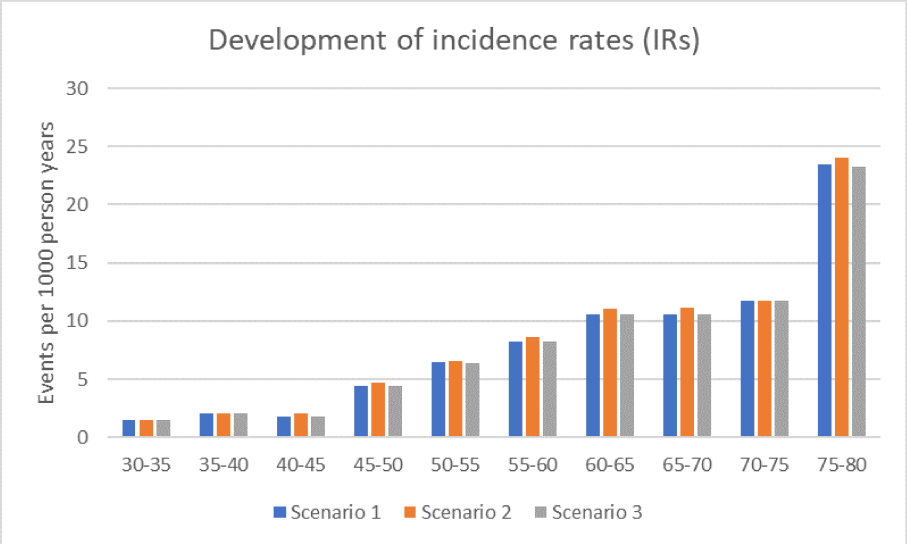


FIGURE 7.2: Incidence rate development between ages 30 and 80 (bins groups of 5 years)

Now that these outputs of selections between the ages 30 and 80 were generated, certain trends can be identified.

It appears that on average, the incidence rates keep increasing with every increase in age bins. Furthermore, IRs range between 1.02 and 24.09, for both the simulation outputs. When looking at the differences among the determined scenarios (IR1, IR2 and IR3 as presented in the tables), most frequent difference is the overall higher IR of scenario 2. As explained in Chapter 6, this is because it is assumed that abnormalities develop constantly in between both points in time. Therefore, more person years are included, but relatively more events as well. When looking at the incidence rates of scenario 1 and 3, these values differ slightly but marginal. According to both the tables, it is most common that the output IRs of scenario 3 are lower, due to a small amount of extra events occurring, with more person years to account for. In this way, the final IR will take a lower value. In the Appendix, figure F.2 and F.3 further zoom on several more simulations of the biggest subject groups (ranging from ages 40 to 70).

7.2 Towards personalized recommendations for a potential revisit

Eventually, we want to work towards a more personalized way of recommending a research moment in future for follow-up. By doing so, a trade-off is considered between the utility of doing a TBS in short notice and on the other hand the risk of developing an abnormality during the period in between these research moments. In order to come to this personalized risk estimation, we match a client's characteristics with a set of data from comparable precedent subjects.

For example, we want to estimate the risk of a 40-year old man. By doing so, potential risk factors as an increase in age or the gender of a subject are equalized, with a more personalized risk as a result.

The difference with the previous paragraph is the overlap in data, and a method to account for uncertainty of the set. If the example of the man of 40 is used initially, the model only selects clients who underwent screening at the exact age of 40. It then compares the related incidence rate with such a selection's somewhere in future. Therefore, when age X is set as input for the model, all clients with ages X-0.5 up to X+0.5 are included. In this way we reduce the risk that results are affected by uncertainty, due to having a too small set of data point for example. On the other hand, clients undergoing screening at an age of 40 or 41, are still seen as sufficiently comparable.

For this specific client of 40 years old, future risk estimations are setup based on this technique by comparing the course of other precedents. The first year of follow-up, is related to the incidence rate of the [LB,UB] combination of 40 up to 42. These subjects can be seen on average as 41, the age of our client by the next year. When

a deviation of ages of 0.5 year is used, the incidence over ages 40 to 41 is taken into account, which is considered as the half year estimation for this same specific client.

In Table 7.1 an example output of this simulation is given. Since it considers a low amount of subjects in each bin, the incidence rates of all three scenarios (IR1, IR2, IR3) all equal either zero or a relatively high value. This is explained by the amount of person years included, and the way of presenting the IR as a rate of events per 1000 person years. In bin [36,37], are included and tracked for a maximum of one single year. If an event occurs in this short period of time, and is recalculated to a rate per 1.000 person years, this single event will be counted as 10 already. When looking at the second part of the table, it can be concluded that with an increased number of subjects in each bin, the IR results in a more stable development. Therefore, same simulations will be performed but with broadened bins of X-1 and X+1. This means a determined time period of two years is used, in which more subjects are included and a better estimation is achieved.

IR1	**IR2**	**IR3**	**LB**	**UB**	**Amount of subjects (n)**
13.83	13.83	13.83	30	31	[redacted]
12.88	12.88	12.88	31	32	[redacted]
11.92	11.92	11.92	32	33	[redacted]
0.00	0.00	0.00	33	34	[redacted]
0.00	0.00	0.00	34	35	[redacted]
0.00	0.00	0.00	35	36	[redacted]
21.65	21.65	21.65	36	37	[redacted]
0.00	3.92	0.00	37	38	[redacted]
0.00	0.00	0.00	38	39	[redacted]
21.17	21.17	21.17	39	40	[redacted]

IR1	**IR2**	**IR3**	**LB**	**UB**	**Amount of subjects (n)**
4.59	4.59	4.59	40	41	[redacted]
4.21	11.18	4.19	41	42	[redacted]
7.34	7.34	7.34	42	43	[redacted]
13.04	13.04	13.04	43	44	[redacted]
8.51	12.22	8.38	44	45	[redacted]
17.37	19.44	17.37	45	46	[redacted]
9.37	11.30	9.36	46	47	[redacted]
22.47	27.67	22.16	47	48	[redacted]
15.57	20.32	15.49	48	49	[redacted]
23.31	25.19	23.30	49	50	[redacted]

TABLE 7.1: Model output of incidence rate development, 1 year increase (ages 30-40 and 40-50)

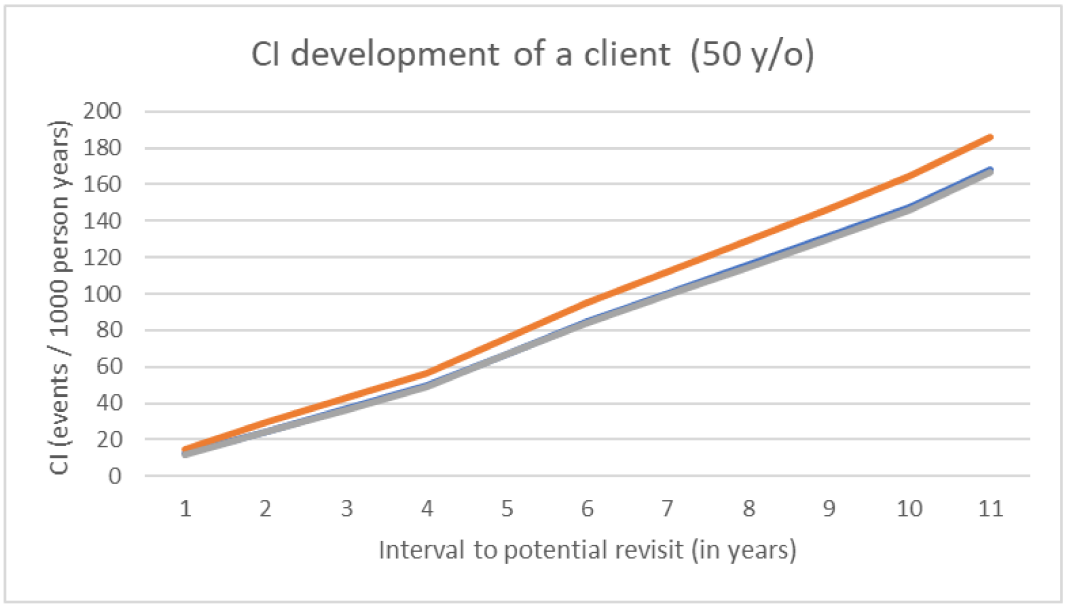


FIGURE 7.3: Cumulative incidence rate and related interval period for revisit with starting age 50

According to the theory of Rothman et al. [67], it appears that when following a cohort through more age classes, for each class a different age specific rate is encountered. Over these periods, the cumulative incidence equals the sum of these IRs, when all are presented within the same unit of events per 1.000 person years. Now, this technique will be applied to each of the subgroup presented in the previous section, starting from ages 30, 40, 50, 60 and 70 respectively. To account for the fact that we are calculating the cumulative incidence yearly, every second year follows Rothman's theory of the next formula:

$$\textit{Cumulative incidence} = IR \cdot T \textit{ (time in years)} \tag{7.1}$$

With this equation, it can be seen that for every second year in an [LB,UB] combination, the same IR is used. Figure 7.3 is a preview of one of these simulation outputs and the three related scenario values. An overview of the summarized results and a comparison between the subgroups is given in figure 7.4. The table outputs and visualizations per age subgroup belonging to this figure are listed in Appendices Figure F.4 and Table F.4 up to Figure F.8 and Table F.8.

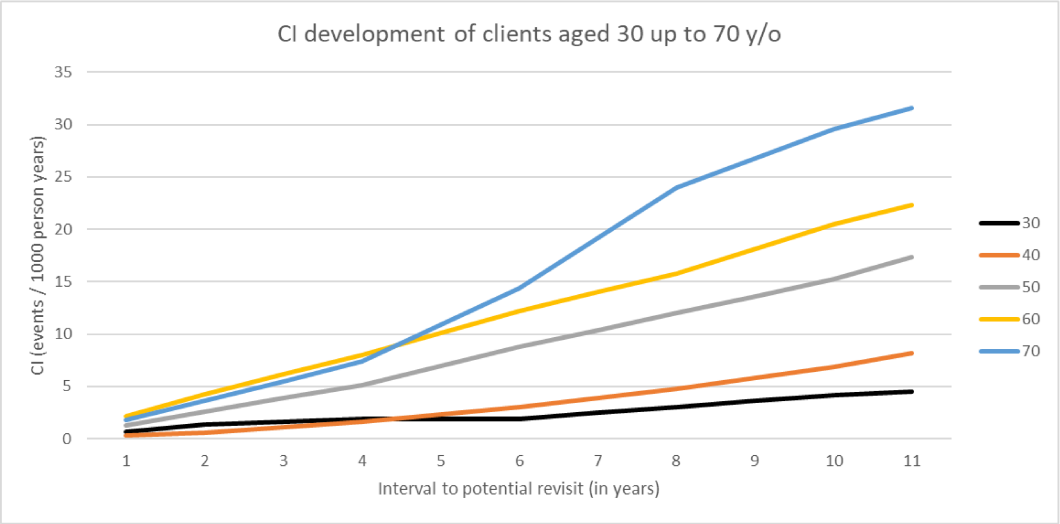


FIGURE 7.4: Overview of CI development as output of different subgroups, comparison between starting ages 30 up to 70

7.3 Unavailable section

This section is not available in public version.

Chapter 8

Conclusion and recommendations

This chapter presents an overview of conclusions that have been made focusing on different parts of this study. First, it will present these main conclusions drawn, by treating all research questions stated in Chapter 1. Then, we will give our main recommendations for Prescan to improve their provision of preventive healthcare towards their clients undergoing screening.

8.1 Conclusions

- *What are current procedures and definitions within the topic of (private) preventive screening?*

Nowadays, being updated about one's health status is an upcoming and a more frequently occurring interest. People's proactive attitude, combined with the fact that the use state-of-the-art imaging techniques in healthcare are becoming more usual and easy in its use, gives an increased interest in the business Prescan is providing; preventive MRI screening. Furthermore, aside from diagnostics in regular healthcare and public screening, this way of individual preventive care has potential to develop towards a standard of early detection -and treatment- of for example chronic or severe diseases.

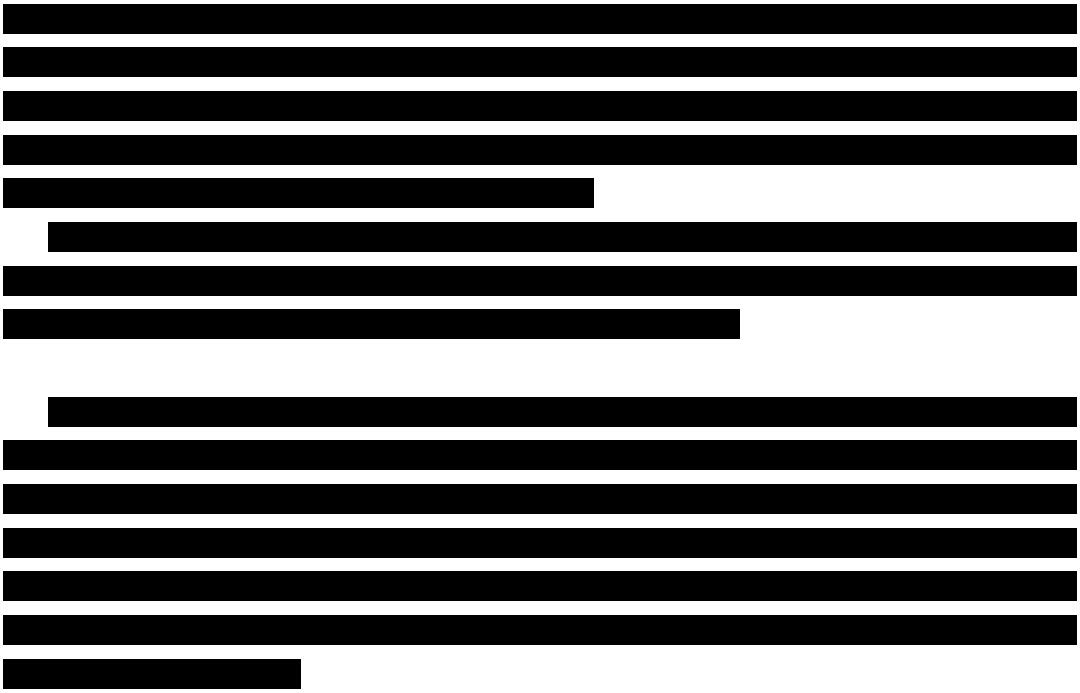
There are many reasons to come up with, when deciding to undergo such a preventive MRI screening. Based on this report's context analysis, it appeared that most frequent considerations made by individual clients doing so, is being reassured about their health status, being able to manage potential risks and avoiding the fear of having regrets later on.

However, when screening a human body, a number of abnormalities can be detected through current imaging techniques. Therefore, the potential risk of a potential clinically relevant finding must be taken into account, and eventual follow-up or recommendations must be provided carefully.

- *What are current strategies and procedures used within Prescan's organization for treating their clients?*

At Prescan, the Total Bodyscan is the most frequent procedure used within their clinics, where MRI scans of several body regions are performed. When an appointment is scheduled, the client visits one of the clinics, and is supported by a mentor throughout the procedure. When the scans are executed by an MRI employee, results will be discussed with the radiologist afterwards. When a clinically relevant finding has been diagnosed, follow-up based on guidelines of regular health care is advised. When no finding has been diagnosed, remarkable deviations or questions are treated. A specific recommendation for a potential revisit does not exist yet, although radiologists do predict what is best for their specific client based on their knowledge and experience.

- *What are possible outcomes and clinically relevant findings of the concerned abdominal MRI within the TBS concept?*



- *Is it feasible to accurately determine the probability of developing and growing a clinically relevant finding based on currently available literature and data, and if so, can this be translated towards a personalized interval?*

After the third chapter's visualizations of all screening moments, useful selections were conducted as shown in Chapter 3. From here, it is concluded that the spread and period in between screening moments of the three findings included are comparable. For clients with a diagnosed aortic finding, average starting age of 65 is higher than starting ages of the pancreatic and renal finding group (60 and 59 respectively). Furthermore, periods of revisiting Prescan after a finding was diagnosed during first screening, are found to be shorter than those of clients with a diagnosis at a later moment.

Although these results give valuable information, it is complicated to appropriately determine development probabilities of clinically relevant findings with only this relatively little data collection of information. Therefore, a literature search was performed, focusing on screening techniques and recommendation for intervals in performing scans. To compare these insights with the existing trends and results from the internal data analysis, this search was performed from the perspective of these selected clinically relevant findings. This gathered literature was used to compare trends and categorize characteristics with the company's existing data. On the contrary, it appeared that little was known when considering frequencies of screening strategies. Aside from public screening and related follow-up for breast and colon cancer, most articles and research setups found are focusing on high-risk population and its cost-effectiveness.

Therefore, in Chapter 6 a model was developed based on articles and definitions focusing on incidence of developing diseases in future, and models to optimize this way of personalized follow-up.

As a result, it can be concluded that recommendations for personalized intervals were achieved. Although it is based on the specific client's age solely, it gives useful insights in the development of their related estimation of developing such a finding. The model is able to develop risk estimation for clients, based on existing data points of comparable clients who underwent MRI screening at Prescan.

Divided into five subgroups of clients, it can be concluded that based on an increased starting age, incidence rates also rise. Based on a predetermined acceptable and preferred risk, the corresponding related interval to revisit for a checkup is recommended, as presented in Table **??**.

- *How can the steps taken to solve the core problem of this research be improved to allow generalization to the other diagnostic tests offered by Prescan?*

In conclusion, the insights resulting from the model calculations can be used as support tool in recommending an interval for clients potentially revisiting Prescan. In this way, it provides with an extra insight for decision making between the radiologist and the clients involved. In order to do so, related assumptions and selections made throughout this study must be taken into account. To come to a generalization of this recommendation, other body regions must be taken into account, and severity of the findings included.

8.2 Recommendations

Since we solely focused on a preselected collection of possible MRI screening outcomes, this is an important characteristic to take into account. Of course, the IR developments found in this report can be used as a support tool for risk estimations. Therefore, these insights can function as support tool in deciding the time horizon

for a potential revisit for screening. By doing so, it must be stated that these developments are solely based on this report's inclusion criteria.

Contribution to Prescan's perspective

With the provided information and insights for clients undergoing preventive screening at Prescan, it can be discussed whether and when to revisit in future. Radiologists may discuss with their clients what an appropriate interval for a revisit in this case would be. This advice is based on the resulting MRI images and related estimations of risks. We would advise to discuss what the maximum probability is to accept and why, in order to get to this recommendation. Therefore, it must also be explained that this report gives clear indication based on a comparable set of former clients, but includes a limited possibility of clinically relevant findings.

The following section treats the contribution to theory, and what possibilities there are to strengthen this recommendation.

Contribution to theory

Considering the context analysis and literature searches performed, this report gives a relevant addition to nowadays interests in medical research. Since not all data from the beginning of Prescan's existence was available, a predetermined selection has been gathered. Because of this deficiency, it solely focuses on a small part of the whole set of possible relevant outcomes within MRI screening. However, it does provide useful insights in estimating client's risk estimation for developing potentially serious findings in future. By doing so, this report shows a way of retrospectively using company's existing data based on theoretical and scientific definitions and models. For a more optimized and sharper individual recommendation for a preventive revisit, it is recommended to gather more data and (behavioral) risk factors.

Chapter 9

Discussion

In the first section of this chapter, limitations of this study will be highlighted per category. Second, suggestions for further research are discussed in Section 9.2.

9.1 Limitations

Use of definitions

The definition of a clinically relevant finding used, requires a clear explanation. When looking at a relatively harmless abnormality, this does not indicate an immediate threat at moment of screening. However, there is a possibility it will further develop to a potential health threat in future. On the contrary, not all deviated values from standards and what seems "normal" must be considered as potential threat. Since we focus on the company's retrospective data, this report therefore uses Prescan's procedures in the definition of a clinically relevant finding. This means that initially all cases were included of which follow-up investigation was advised, to gather all the findings diagnosed through the related MRI screening. For the other group of clients, of which finding had been diagnosed, an MRI screening at a later stage is called a "revisit", instead of a "follow-up" investigation.

Availability of data

The initial purpose of this study was to come up with an optimization of personalized intervals for clients revisiting Prescan. In order to do so, it appeared from several literature studies that data are required of clients with known screening moments, and a sufficient number of findings has been observed. With this information it is known when a client was not identified with a finding, until the screening moment that they were. Then, probabilities of these developments could have been used as input for the intended optimization.

However, as mentioned previously, outcomes of all performed scans in Prescan's history were not registered structurally. Therefore, in this report a database is prepared, in which a preselected group of clients with a diagnosed finding were included. By doing so, results and certain trends give insights for clients and personnel, but must be used carefully when it is applied within a recommendation for

a revisit. Since it includes a small part of potential outcomes of MRI screening, it requires further research and data gathering to improve external validity and generalization.

When discussing the amount of gathered data, in this report we were able to expose individual screening moments for individuals who visited Prescan in between October 2016 and May 2018. For these clients it was possible to retrospectively gather all related MRI scans performed. This resulted in 414 clients with a diagnosed finding. It therefore seems interesting to compare this data with other possible outcomes in the future, or perform same procedures in a later stage so that more comparable data is available.

[redacted]

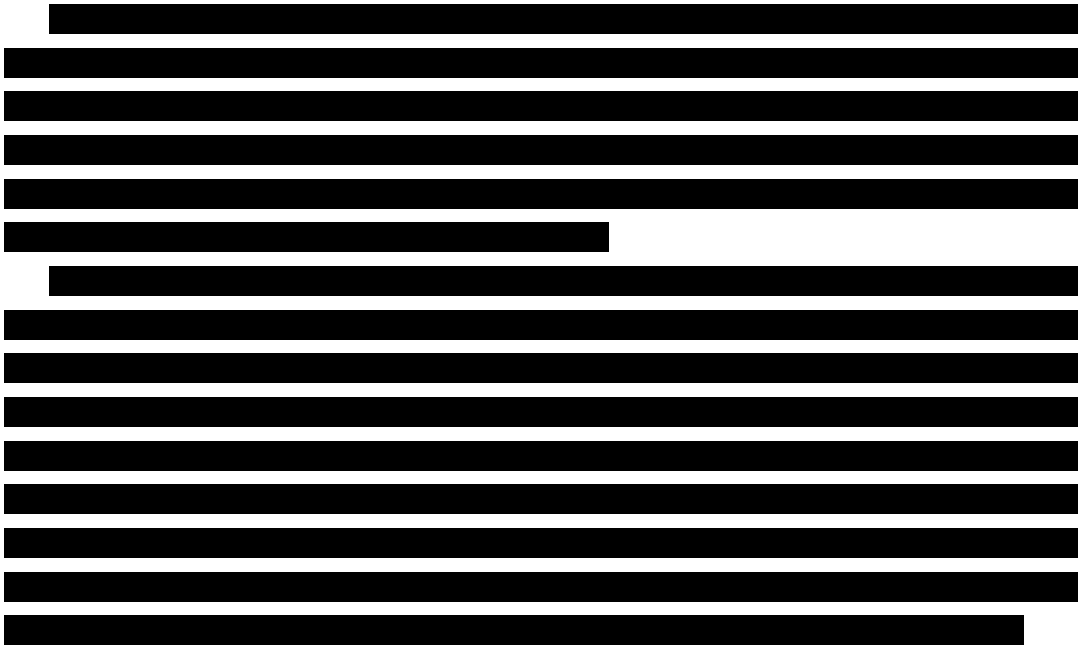
Research population

If a comparison with existing scientific research is made, it must be considered to which extent this report's research population can be generalized. This consideration also holds when recommending an individual client, since the results are based on trends in former clients forming the research population.

When considering generalization of the results, because this proactive behavior of preventive screening reflects their daily lifestyle in a positive way, this study's research population therefore cannot be fully compared to for example the general Dutch population. In order to do so, individual characteristics and behavior, genetic factors or for example family disease history can be added to strengthen this comparison.

Results

Disease development Before discussing the results as output of the model for calculating incidence rates, insights gathered throughout the report are worth discussing. When looking at the data collection of the 414 aortic (n=36), pancreatic (n=217) and renal findings (n=161), 118 of them were revisiting clients. [redacted]



Subgroup selection In the presentation of the results as outcome of the model, selection of subgroups were made, each with a different starting age. In this way, the development of incidence rates are given per starting age of each group. To come up with a useful overview, ages are determined to start every decade. As a result, development in this risk estimation is given based on subgroups for 40-, 50-, 60- and 70-year old individuals. When it is preferred to gain knowledge for in-between results, this can be achieved by resetting and running the model with different input values.

Scenario setup When looking at the setup of the different scenarios within the model in Section 6.3, it is stated that the first scenario reflects the usual definition of an incidence rate. This indicates that clients with a diagnosed finding are excluded from calculation after the corresponding moment of screening. But, moments of screening that are outside of the determined [LB,UB]-bin, are fully excluded from the last moment in this specific bin onward.

Therefore, second and third scenario are used to visualize the effect of this exclusion. Screening moments taking place later than current [LB,UB] combination, are included in the calculation of incidence rates partly. In scenario 3 it is assumed that if a finding will occur it develops within one year after the negative screening moment. However, it is disease dependent whether this one year horizon is appropriate, which makes this assumption arbitrary. In the end one can not know for sure whether an individual will develop a specific disease or not and if so, when this will occur.

9.2 Further research

Practical recommendations

First, as previously mentioned throughout the report, it must be stated that the results are calculated by using retrospective data with multiple subgroups of Prescan's former clients. Therefore, it is relevant to mention this when these insights are used towards an individual recommendation.

Gathering structured data To expand the domain of current results, further research can contribute to improved risk estimations from both a theoretical and practical perspective.

First of all, based on the related age, recommendations are given following from the calculated risk estimations in developing a potential clinically relevant finding. In order to come to even more personalized advice, individual characteristics can be included in the data gathering, such as gender. Since the current set of positive findings is considered small already, adding these characteristics will result into even smaller subsets. Therefore these analyses are not within the scope of this report. When it is preferred to get these insights, the amount of available data must be increased.

Furthermore, information about certain behavioral risk factors like having a history with smoking, use of alcohol or obesity can add value to future research as well. This can be achieved by the company itself by registering these data more structurally. In addition, a focus on gender seems valuable as well. In line with studies focusing on a personalisation of disease follow-up (or revisiting strategy), it is recommended to track down the size, behavior and stages of clinical findings to show development over time. In this way, it enables to predict general risks in developing thesee findings [60].

Results validation Considering the validation of the results, it is recommended for Prescan to keep registering outcomes of performed MRI scans in a structured way, based on procedures used in regular healthcare. Since these guidelines already include automated thresholds for severity of diseases, this can validate the comparisons made in Chapter 5 between empirical findings and trends in medical research.

As an addition, when recalling the definition used for clinically relevant findings throughout this report: further research can also focus on the difference between an average advised follow-up investigation, or follow-up within three months as considered in the described study in Chapter 2. When excluding all follow-up periods between moment of diagnosis and the screening moment after, current results can be validated by comparing with this new selection.

Considering the censoring principle, it can also be of significant value to keep track of clients' health status who visited for screening. Specifically after screening this can be interesting, when one has been treated within regular healthcare for example.

To improve validity and confidence of the achieved recommendations, the same model setup can be used for more data, or information of clients with other clinically relevant findings. Another possibility is to reproduce the current dataset by random sampling and replacement, which is called bootstrapping. This technique allows the assignment of accuracy measures within the results.

Scientific recommendations

Towards optimization Furthermore, insights are presented based on the outcome of an MRI screening being dichotomous. When more data is produced, it will become of more added value to zoom in on development and severity of diagnosed findings. Looking at the input of the Markov model produced by Otten et al. [65], this transition probability appears to be the core input of the model. In this report, we do not have the required information to implement an optimization like the Partially Observable Markov Decision Process (POMDP) described. When looking at the determination of this required transition probability, regression analyses are performed over large datasets of cancer patients (50.000+) and their corresponding moments of treatment and risk factors [60]. If Prescan is willing to record all these required data from their clients, such a model can be used to achieve a next step in the optimization of their client's personalized revisiting strategy.

When this would be applied, it becomes of more added value to apply statistical techniques for achieving a more solid recommendation per individual, with the use of more data and characteristics.

Bibliography

[1] RIVM, *Bevolkingsonderzoeken en screeningen - RIVM*, 2012. [Online]. Available: https://www.rivm.nl/Onderwerpen/B/Bevolkingsonderzoeken_en_screeningen/Screening_de_theorie/Wet_en_regelgeving.

[2] KNMG, "Preventief Medisch Onderzoek", 2015. [Online]. Available: https://www.knmg.nl/advies-richtlijnen/dossiers/preventief-medisch-onderzoek.htmhttp://www.knmg.nl/Nieuws/Overzicht-nieuws/Nieuwsbericht/149538/KNMG-informeert-consument-over-risicos-gezondheidschecks.htm.

[3] T. Boer, S. Einerhand, J. de Haas-de Vries, and M. van Rijswijk, *Komt een test bij de dokter*. 2018, p. 132.

[4] Prescan Nederland, *De Total Bodyscan | Prescan® - De beste zorg is uit voorzorg*, 2018. [Online]. Available: https://www.prescan.nl/.

[5] RTV Oost, *Prescan mag MRI-scans uitvoeren*, 2009. [Online]. Available: http://www.rtvoost.nl/nieuws/default.aspx?nid=105118&rubriek=2.

[6] M. Hoffmann and R. Schmücker, "The Ethics of Incidental Findings in Population-Based MRI Research", in *Whole-body MRI Screening*, R. Puls and N. Hosten, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 1–19, ISBN: 978-3-642-55201-4. DOI: 10.1007/978-3-642-55201-4{_}1. [Online]. Available: https://doi.org/10.1007/978-3-642-55201-4_1.

[7] ONVZ, *Fit blijven - ONVZ Zorgverzekeraar*, 2018. [Online]. Available: https://www.onvz.nl/zelf-regelen/de-zorgconsulent/fit-blijven.

[8] Prescan, *Prescan® Total Bodyscan*, 2018. [Online]. Available: https://www.prescan.nl/total-body-scan/.

[9] E. I. Schippers, *Nader standpunt health checks (kamerbrief)*, 2016.

[10] KNMG, "Autorisatie, Beheer, Implementatie, Evaluatie Multidisciplinaire Richtlijn Preventief Medisch Onderzoek", KNMG, Utrecht, Tech. Rep., 2017, p. 51.

[11] G. Klaß and G. Kindervater, *Interview Rheine Mathias Spital*, Rheine, 2018.

[12] H. H. Quick, "Technical Prerequisites for Whole-Body MRI Screening", in *Whole-body MRI Screening*, R. Puls and N. Hosten, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 29–44, ISBN: 978-3-642-55201-4. DOI: 10.1007/978-3-642-55201-4{_}3. [Online]. Available: https://doi.org/10.1007/978-3-642-55201-4_3.

[13] RadiologyInfo, *Body MRI - magnetic resonance imaging of the chest, abdomen and pelvis*, 2016. [Online]. Available: https://www.radiologyinfo.org/en/info.cfm?pg=bodymr.

[14] W. Oldendorf, "Advantages and Disadvantages of MRI", *Life Science Journal*, vol. 10, no. SUPPL.4, pp. 458–462, 2013, ISSN: 0276-3893. DOI: 10.1007/978-1-4613-2081-4{_}9. [Online]. Available: http://link.springer.com/10.1007/978-1-4613-2081-4_9.

[15] C. A. Yi, K. M. Shin, K. S. Lee, B.-T. Kim, H. Kim, O. J. Kwon, J. Y. Choi, and M. J. Chung, "Non–Small Cell Lung Cancer Staging: Efficacy Comparison of Integrated PET/CT versus 3.0-T Whole-Body MR Imaging", *Radiology*, vol. 248, no. 2, pp. 632–642, Aug. 2008, ISSN: 0033-8419. DOI: 10.1148/radiol.2482071822. [Online]. Available: http://pubs.rsna.org/doi/10.1148/radiol.2482071822.

[16] S. C. Ladd, "Whole-body mri as a screening tool?", *European journal of radiology*, vol. 70, no. 3, pp. 452–462, 2009.

[17] McGraw-Hill, "Incidentaloma", in *Concise Dictionary of Modern Medicine*, 2002. [Online]. Available: https://medical-dictionary.thefreedictionary.com/incidentaloma.

[18] S. M. Wolf, F. P. Lawrenz, C. A. Nelson, J. P. Kahn, M. K. Cho, E. W. Clayton, J. G. Fletcher, M. K. Georgieff, D. Hammerschmidt, K. Hudson, J. Illes, V. Kapur, M. A. Keane, B. A. Koenig, B. S. LeRoy, E. G. McFarland, J. Paradise, L. S. Parker, S. F. Terry, B. Van Ness, and B. S. Wilfond, "Managing Incidental Findings in Human Subjects Research: Analysis and Recommendations", *The Journal of Law, Medicine & Ethics*, vol. 36, no. 2, pp. 219–248, Jun. 2008, ISSN: 1073-1105. DOI: 10.1111/j.1748-720X.2008.00266.x. [Online]. Available: http://journals.sagepub.com/doi/10.1111/j.1748-720X.2008.00266.x.

[19] A. E. Board, *Abdominal MRI scan*, 2012. [Online]. Available: https://www.uclahealth.org/endocrine-center/abdominal-mri-scanhttp://www.ncbi.nlm.nih.gov/books/PMH0004253/.

[20] WebMD, *Magnetic Resonance Imaging (MRI) For Brain, Heart, Breast, Abdomen*, 2018. [Online]. Available: https://www.webmd.com/a-to-z-guides/when-do-i-need-an-mri#3.

[21] D. O'Connor, S. Green, and J. P. Higgins, *Defining the Review Question and Developing Criteria for Including Studies*. 2008, pp. 81–94, ISBN: 9780470699515. DOI: 10.1002/9780470712184.ch5.

[22] A. Booth, "Clear and present questions: Formulating questions for evidence based practice", *Library Hi Tech*, vol. 24, no. 3, pp. 355–368, 2006, ISSN: 07378831. DOI: 10.1108/07378830610692127.

[23] P. F. V. Gils, "Kosteneffectiviteit van preventie", *RIVM*, 2008.

[24] C. Longo and G. R. Upchurch, "Abdominal aortic aneurysm screening: Recommendations and controversies", *Vascular and Endovascular Surgery*, vol. 39, no. 3, pp. 213–219, 2005, ISSN: 15385744. DOI: 10.1177/153857440503900301.

[25] A. Sisó-Almirall, B. Kostov, M. N. González, D. C. Salami, A. P. Jiménez, R. G. Soleâ, C. B. Saumell, L. D. Bach, M. V. Martí, L. G. D. Paz, R. R. Riera, V. R. Alonso, N. Acar-Denizli, M. F. Almacellas, M. Ramos-Casals, and J. B. Àreu, "Abdominal aortic aneurysm screening program using hand-held ultrasound in primary healthcare", *PLoS ONE*, vol. 12, no. 4, pp. 1–16, 2017, ISSN: 19326203. DOI: 10.1371/journal.pone.0176877.

[26] D. P. J. Howard, A. Banerjee, J. F. Fairhead, A. Handa, L. E. Silver, P. M. Rothwell, and Oxford Vascular Study, "Population-Based Study of Incidence of Acute Abdominal Aortic Aneurysms With Projected Impact of Screening Strategy.", *Journal of the American Heart Association*, vol. 4, no. 8, e001926, 2015, ISSN: 2047-9980. DOI: 10.1161/JAHA.115.001926. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/26289347%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4599457.

[27] S. Svensjö, M. Björck, and A. Wanhainen, "Update on screening for abdominal aortic aneurysm: A topical review", *European Journal of Vascular and Endovascular Surgery*, vol. 48, no. 6, pp. 659–667, 2014, ISSN: 15322165. DOI: 10.1016/j.ejvs.2014.08.029. [Online]. Available: http://dx.doi.org/10.1016/j.ejvs.2014.08.029.

[28] L. Meecham, J. Jacomelli, A. D. Pherwani, and J. Earnshaw, "Editor's Choice – Self-referral to the NHS Abdominal Aortic Aneurysm Screening Programme", *European Journal of Vascular and Endovascular Surgery*, vol. 52, no. 3, pp. 317–321, 2016, ISSN: 15322165. DOI: 10.1016/j.ejvs.2016.04.002. [Online]. Available: http://dx.doi.org/10.1016/j.ejvs.2016.04.002.

[29] R. A. P. Scott, S. G. Bridgewater, and H. A. Ashton, "Randomized clinical trial of screening for abdominal aortic aneurysm in women", *British Journal of Surgery*, vol. 89, no. 3, pp. 283–285, Mar. 2002, ISSN: 0007-1323. DOI: 10.1046/j.0007-1323.2001.02014.x. [Online]. Available: http://www.blackwell-synergy.com/links/doi/10.1046/j.0007-1323.2001.02014.x.

[30] J. W. Hallett, "Management of Abdominal Aortic Aneurysms", *Mayo Clinic Proceedings*, vol. 75, no. 4, pp. 395–399, Apr. 2000, ISSN: 00256196. DOI: 10.4065/75.4.395. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0025619611640366.

[31] R. A. Benson, R. Poole, S. Murray, P. Moxey, and I. M. Loftus, "Screening results from a large United Kingdom abdominal aortic aneurysm screening center in the context of optimizing United Kingdom National Abdominal Aortic Aneurysm Screening Programme protocols", *Journal of Vascular Surgery*, vol. 63, no. 2, pp. 301–304, 2016, ISSN: 10976809. DOI: 10.1016/j.jvs.2015.08.091. [Online]. Available: http://dx.doi.org/10.1016/j.jvs.2015.08.091.

[32] B. G. DeRubertis, S. M. Trocciola, E. J. Ryer, F. M. Pieracci, J. F. McKinsey, P. L. Faries, and K. C. Kent, "Abdominal aortic aneurysm in women: Prevalence, risk factors, and implications for screening", *Journal of Vascular Surgery*, vol. 46, no. 4, pp. 16–20, 2007, ISSN: 07415214. DOI: 10.1016/j.jvs.2007.06.024.

[33] A. Jawien, B. Formankiewicz, T. Derezinski, A. Migdalski, P. Brazis, and L. Woda, "Abdominal aortic aneurysm screening program in Poland", *Gefässchirurgie*, vol. 19, no. 6, pp. 545–548, 2014, ISSN: 0948-7034. DOI: 10.1007/s00772-014-1348-4. [Online]. Available: http://link.springer.com/10.1007/s00772-014-1348-4.

[34] N. Olchanski, A. Winn, J. T. Cohen, and P. J. Neumann, "Abdominal aortic aneurysm screening: How many life years lost from underuse of the medicare screening benefit?", *Journal of General Internal Medicine*, vol. 29, no. 8, pp. 1155–1161, 2014, ISSN: 15251497. DOI: 10.1007/s11606-014-2831-z.

[35] M. X. Lee, M. W. Saif, and M. New, "Screening for Early Pancreatic Ductal Adenocarcinoma : An Urgent Call !", vol. 10, no. 2, pp. 104–108, 2009.

[36] M.-C. Chang, "Screening and early detection of pancreatic cancer in high risk population", *World Journal of Gastroenterology*, vol. 20, no. 9, p. 2358, 2014, ISSN: 1007-9327. DOI: 10.3748/wjg.v20.i9.2358. [Online]. Available: http://www.wjgnet.com/1007-9327/full/v20/i9/2358.htm.

[37] F. O. R. Screening, "Guidelines for Screening", vol. 194, no. 1, pp. 1–6, 2015, ISSN: 13697137. DOI: 10.3109/13697137.2011.543017.

[38] C. Hur, A. C. Tramontano, E. C. Dowling, A Brooks, A. Jeon, W. R. Brugge, and G. S. Gazelle, "on Size : Positive Implications for Future Targeted Screening", vol. 45, no. 7, pp. 1062–1066, 2017. DOI: 10.1097/MPA.0000000000000587.Early.

[39] R. M. Gore, D. R. Wenzke, K. H. Thakrar, G. M. Newmark, U. K. Mehta, and J. W. Berlin, "The incidental cystic pancreas mass: A practical approach", *Cancer Imaging*, vol. 12, no. 2, pp. 414–421, 2012, ISSN: 14707330. DOI: 10.1102/1470-7330.2012.9054.

[40] A. E. Becker, Y. G. Hernandez, H. Frucht, and A. L. Lucas, "Pancreatic ductal adenocarcinoma: Risk factors, screening, and early detection", *World Journal of Gastroenterology*, vol. 20, no. 32, pp. 11 182–11 198, 2014, ISSN: 22192840. DOI: 10.3748/wjg.v20.i32.11182.

[41] R. Pezzilli, D. Fabbri, and A. Imbrogno, "Pancreatic ductal adenocarcinoma screening new perspectives.", *World journal of gastroenterology : WJG*, vol. 18, no. 36, pp. 4973–7, 2012, ISSN: 1007-9327. DOI: 10.3748/wjg.v18.i36.4973. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/23407480.

[42] K. E. Poruk, M. A. Firpo, D. G. Adler, and Mulvihill Sean J, "NIH Public Access", *Ann Surg*, vol. 257, no. 1, pp. 17–26, 2013. DOI: 10.1097/SLA.0b013e31825ffbfb.Screening.

[43] K. Fatema, Z. Abedin, A. Mansur, F. Rahman, T. Khatun, N. Sumi, K. Kobura, S. Akter, and L. Ali, "Screening for chronic kidney diseases among an adult population.", *Saudi journal of kidney diseases and transplantation : an official publication of the Saudi Center for Organ Transplantation, Saudi Arabia*, vol. 24, no. 3, pp. 534–41, 2013, ISSN: 1319-2442. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/23640626.

[44] C.-C. Chang, J.-Y. Kuo, W.-L. Chan, K.-K. Chen, and L. S. Chang, "Prevalence and clinical characteristics of simple renal cyst.", *Journal of the Chinese Medical Association : JCMA*, vol. 70, no. 11, pp. 486–491, 2007, ISSN: 1726-4901. DOI: 10.1016/S1726-4901(08)70046-7.

[45] B. S. Malaeb, D. J. Martin, F. N. Littooy, Y. Lotan, W. B. Waters, R. C. Flanigan, and K. S. Koeneman, "The utility of screening renal ultrasonography: Identifying renal cell carcinoma in an elderly asymptomatic population", *BJU International*, vol. 95, no. 7, pp. 977–981, 2005, ISSN: 14644096. DOI: 10.1111/j.1464-410X.2005.05451.x.

[46] A. Sohaib, "Incidental solid cystic renal lesion", *Cancer Imaging*, vol. 12, no. 2, pp. 385–386, 2012, ISSN: 14707330. DOI: 10.1102/1470-7330.2012.9034.

[47] S. Paulo, M. Francisco, D. Oglio, M. Srougi, and P. D. Gonçalves, "Incidental and symptomatic renal tumors :", vol. 120, no. 6, pp. 165–169, 2002.

[48] F. Levi, J. Ferlay, C. Galeone, F. Lucchini, E. Negri, P. Boyle, and C. La Vecchia, "The changing pattern of kidney cancer incidence and mortality in Europe", *BJU International*, vol. 101, no. 8, pp. 949–958, 2008, ISSN: 14644096. DOI: 10.1111/j.1464-410X.2008.07451.x.

[49] J. J. Patard, "Incidental renal tumours", *Current Opinion in Urology*, vol. 19, no. 5, pp. 454–458, 2009, ISSN: 09630643. DOI: 10.1097/MOU.0b013e32832f0ccd.

[50] A. B. Wilmink, C. S. Hubbard, N. E. Day, and C. R. Quick, "The incidence of small abdominal aortic aneurysms and the change in normal infrarenal aortic diameter: Implications for screening", *European Journal of Vascular and Endovascular Surgery*, vol. 21, no. 2, pp. 165–170, 2001, ISSN: 10785884. DOI: 10.1053/ejvs.2000.1285.

[51] Y. Takeda, A. Saiura, Y. Takahashi, Y. Inoue, T. Ishizawa, Y. Mise, M. Matsumura, H. Ichida, R. Matsuki, M. Tanaka, and H. Ito, "Asymptomatic Pancreatic Cancer: Does Incidental Detection Impact Long-Term Outcomes?", *Journal of Gastrointestinal Surgery*, vol. 21, no. 8, pp. 1287–1295, 2017, ISSN: 1091255X. DOI: 10.1007/s11605-017-3421-2.

[52] K. Rungsinaporn and T. Phaisakamas, "Frequency of abnormalities detected by upper abdominal ultrasound", *Journal of the Medical Association of Thailand*, vol. 91, no. 7, pp. 1072–1075, 2008, ISSN: 01252208.

[53] M Oshibuchi, F Nishi, M Sato, H Ohtake, and K Okuda, "Frequency of abnormalities detected by abdominal ultrasound among Japanese adults.", *Journal of gastroenterology and hepatology*, vol. 6, no. 2, pp. 165–8, ISSN: 0815-9319. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/1912425.

[54] L. L. Berland, S. G. Silverman, R. M. Gore, W. W. Mayo-Smith, A. J. Megibow, J. Yee, J. A. Brink, M. E. Baker, M. P. Federle, W. D. Foley, I. R. Francis, B. R. Herts, G. M. Israel, G. Krinsky, J. F. Platt, W. P. Shuman, and A. J. Taylor, "Managing incidental findings on abdominal CT: White paper of the ACR incidental findings committee", *Journal of the American College of Radiology*, vol. 7, no. 10, pp. 754–773, 2010, ISSN: 15461440. DOI: 10.1016/j.jacr.2010.06.013. [Online]. Available: http://dx.doi.org/10.1016/j.jacr.2010.06.013.

[55] J. M. G. Wilson and G. Jungner, "Principles and Practice of Screening for Disease.", *Archives of Internal Medicine*, vol. 123, no. 3, p. 349, 1969, ISSN: 0003-9926. DOI: 10.1001/archinte.1969.00300130131020. [Online]. Available: http://archinte.jamanetwork.com/article.aspx?doi=10.1001/archinte.1969.00300130131020.

[56] I. M. Burger and N. E. Kass, "Screening in the dark: Ethical considerations of providing screening tests to individuals when evidence is insufficient to support screening populations", *The American Journal of Bioethics*, vol. 9, no. 4, pp. 3–14, 2009.

[57] C. Wohlin, "Guidelines for snowballing in systematic literature studies and a replication in software engineering", in *Proceedings of the 18th international conference on evaluation and assessment in software engineering*, ACM, 2014, p. 38.

[58] L. Dong, S. Y. Hu, Q. Zhang, R. M. Feng, L. Zhang, X. L. Zhao, X. Zhang, Y. L. Qiao, and F. H. Zhao, "Risk prediction of cervical cancer and precancers by type-specific human papillomavirus: Evidence from a population-based cohort study in China", *Cancer Prevention Research*, vol. 10, no. 12, pp. 745–751, 2017, ISSN: 19406215. DOI: 10.1158/1940-6207.CAPR-17-0088.

[59] J. M. Hostetter, J. J. Morrison, M. Morris, J. Jeudy, K. C. Wang, and E. Siegel, "Personalizing lung cancer risk prediction and imaging follow-up recommendations using the National Lung Screening Trial dataset", *Journal of the American Medical Informatics Association*, vol. 24, no. 6, pp. 1046–1051, 2017. DOI: 10.1093/jamia/ocx012. [Online]. Available: http://dx.doi.org/10.1093/jamia/ocx012.

[60] A. Witteveen, J. W. M. Otten, I. M. H. Vliegen, S. Siesling, J. B. Timmer, and M. J. IJzerman, "Risk-based breast cancer follow-up stratified by age.", eng, *Cancer medicine*, vol. 7, no. 10, pp. 5291–5298, 2018, ISSN: 2045-7634 (Electronic). DOI: 10.1002/cam4.1760.

[61] T. M. Ripping, R. A. Hubbard, J. D. M. Otten, G. J. den Heeten, A. L. M. Verbeek, and M. J. M. Broeders, "Towards personalized screening: Cumulative risk of breast cancer screening outcomes in women with and without a first-degree relative with a history of breast cancer.", eng, *International journal of cancer*, vol. 138, no. 7, pp. 1619–1625, 2016, ISSN: 1097-0215 (Electronic). DOI: 10.1002/ijc.29912.

[62] C. M. Tammemagi, P. F. Pinsky, N. E. Caporaso, P. A. Kvale, W. G. Hocking, T. R. Church, T. L. Riley, J. Commins, M. M. Oken, C. D. Berg, and P. C. Prorok, "Lung cancer risk prediction: Prostate, Lung, Colorectal And Ovarian Cancer Screening Trial models and validation.", eng, *Journal of the National Cancer Institute*, vol. 103, no. 13, pp. 1058–1068, 2011, ISSN: 1460-2105 (Electronic). DOI: 10.1093/jnci/djr173.

[63] T. Ayer, O. Alagoz, and N. K. Stout, "OR Forum—A POMDP Approach to Personalize Mammography Screening Decisions", *Operations Research*, vol. 60, no. 5, pp. 1019–1034, 2012, ISSN: 0030-364X. DOI: 10.1287/opre.1110.1019. [Online]. Available: http://pubsonline.informs.org/doi/abs/10.1287/opre.1110.1019.

[64] J. Zhang, B. T. Denton, H. Balasubramanian, N. D. Shah, and B. A. Inman, "Optimization of PSA screening policies: A comparison of the patient and societal perspectives", *Medical Decision Making*, vol. 32, no. 2, pp. 337–349, 2012, ISSN: 0272989X. DOI: 10.1177/0272989X11416513.

[65] J. Otten, J. Timmer, and A. Witteveen, *Stratified breast cancer follow-up using a continuous state partially observable Markov decision process*, English, ser. TW-Memoranda 2064. Department of Applied Mathematics, University of Twente, Apr. 2018.

[66] B MacMahon and T. F. Pugh, *Epidemiology: Principles and Methods*. Little, Brown, 1970. [Online]. Available: https://books.google.nl/books?id=MLIXAAAAIAAJ.

[67] K. J. Rothman, S. Greenland, T. L. Lash, *et al.*, "Modern epidemiology", 2008.

[68] S. Kotz and J. R. van Dorp, *Beyond Beta*. WORLD SCIENTIFIC, 2004, ISBN: 978-981-256-115-2. DOI: 10.1142/5720. [Online]. Available: https://www.worldscientific.com/worldscibooks/10.1142/5720.

[69] S. Robinson, *Simulation : the practice of model development and use LK - https://ut.on.worldcat.org/oclc/8844*
Engels, Second edition. Houndmills, Basingstoke, Hampshire, UK ; SE - xxii, 367 pages ; 26 cm: Palgrave Macmillan, 2014, ISBN: 9781137328021 1137328029.

Appendix A

Confidential Appendix

This section is not available in the public version.

Appendix B

Data and literature gathering

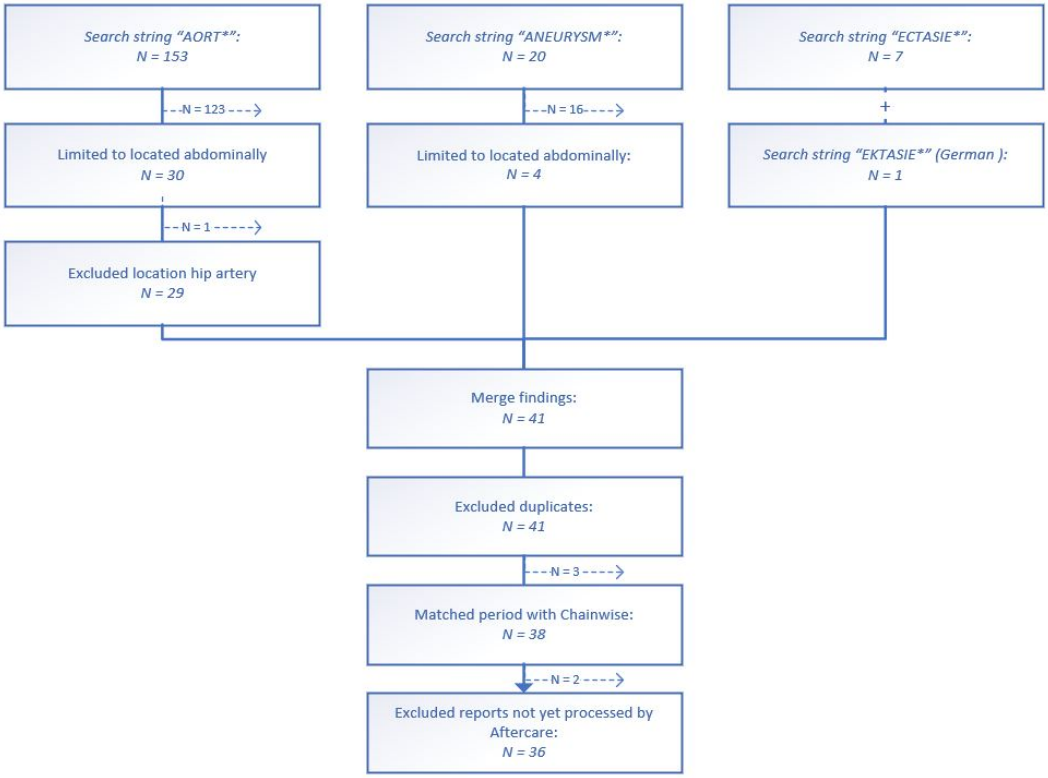


FIGURE B.1: An overview of identified abnormalities within the abdominal aorta

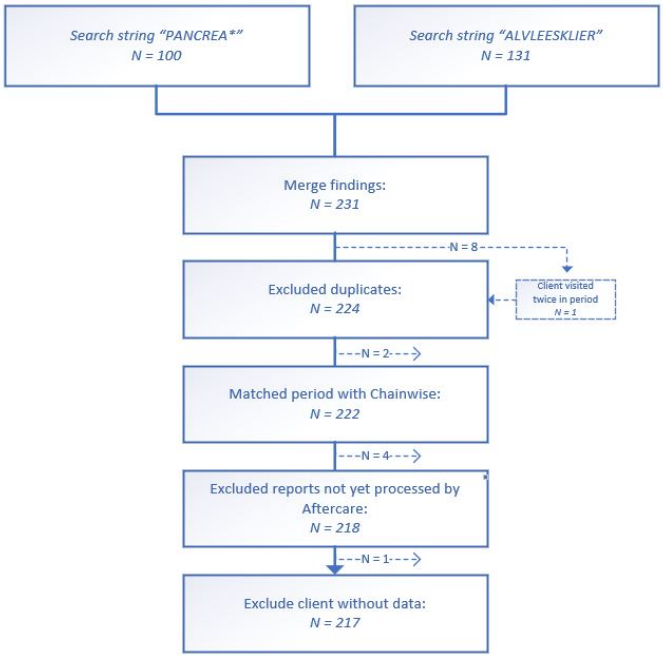


FIGURE B.2: An overview of identified abnormalities within the pancreas

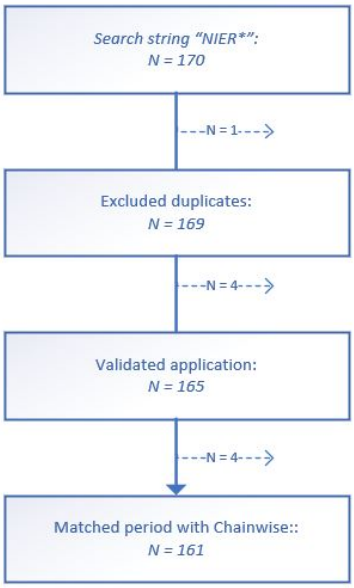


FIGURE B.3: An overview of identified abnormalities within the kidneys

Appendix C

Abnormality selection flowcharts

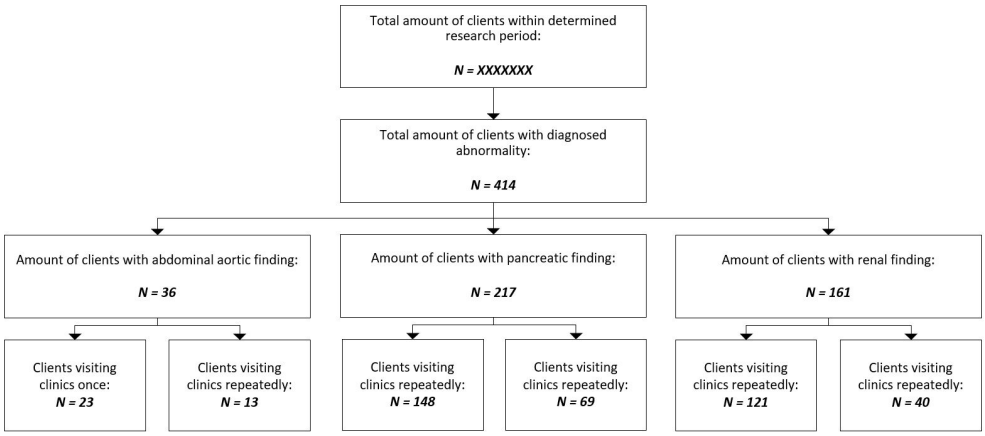


FIGURE C.1: An overview of identified abnormalities within the abdominal aorta, pancreas and kidneys

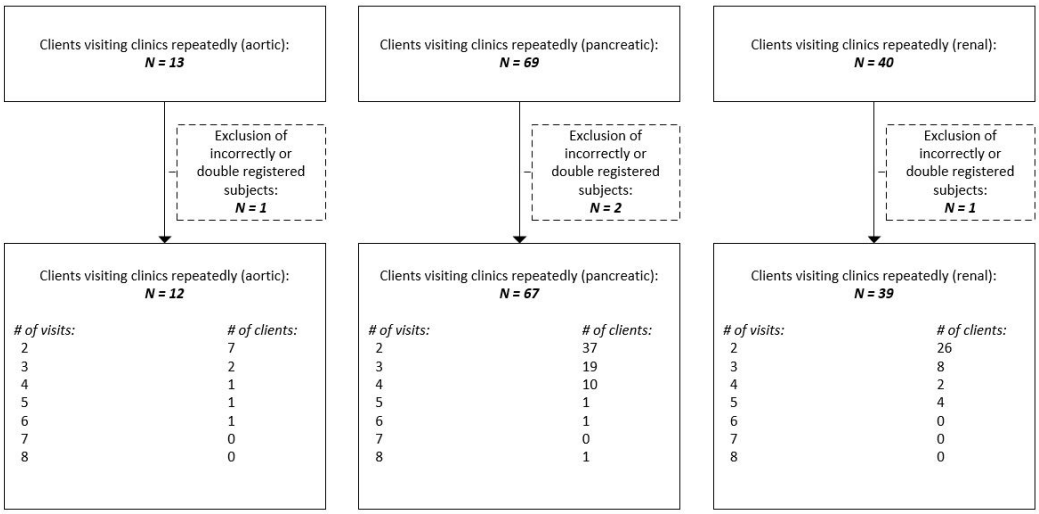


FIGURE C.2: An overview of identified abnormalities within the abdominal aorta, pancreas and kidneys

Appendix D

Flowcharts of literature searches

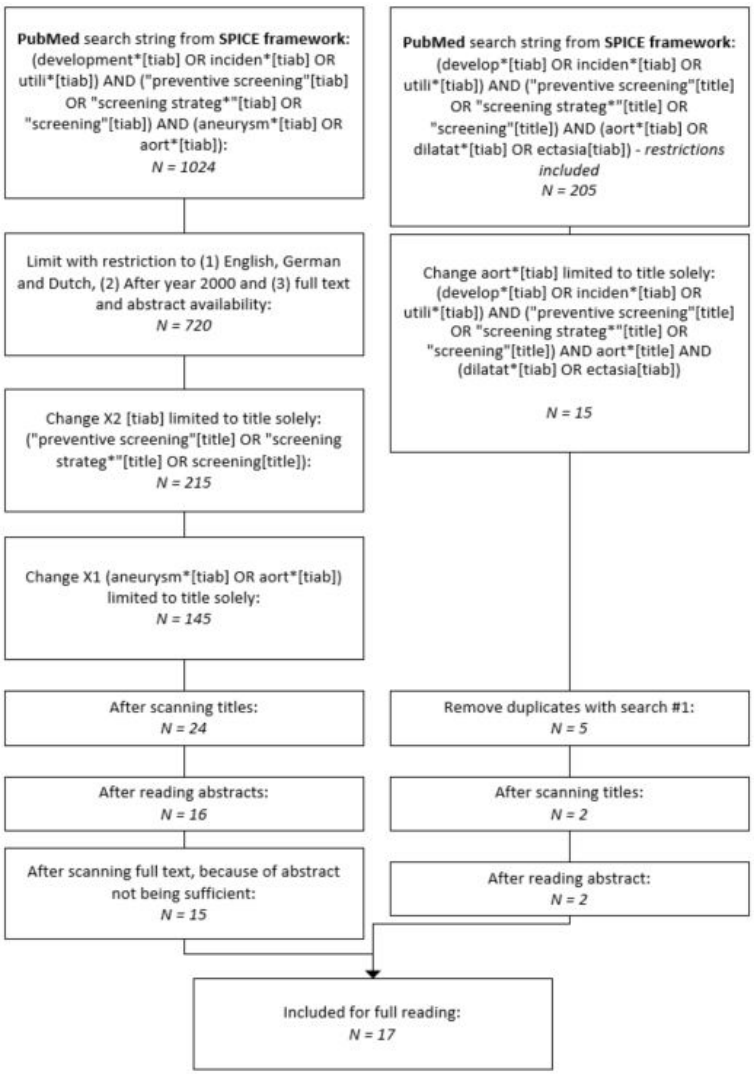


FIGURE D.1: Article flow of literature search (Aorta)

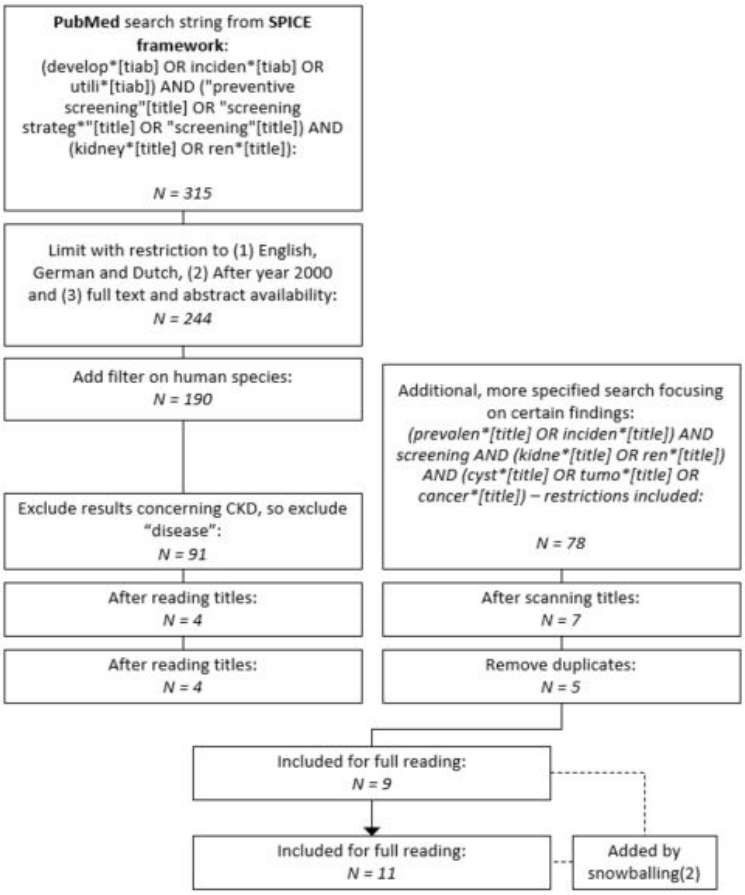


FIGURE D.2: Article flow of literature search (Pancreas)

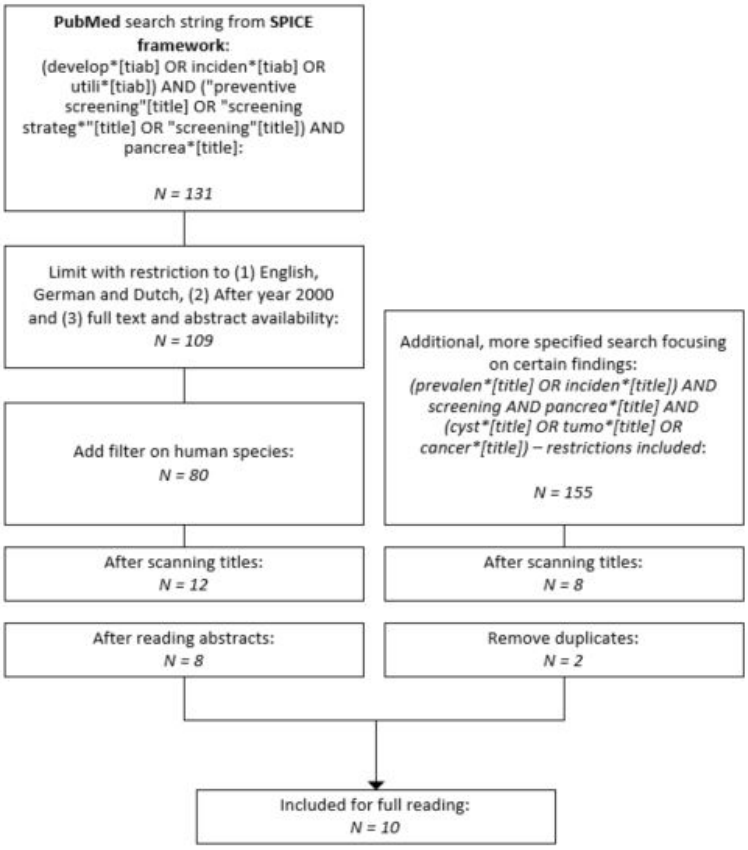


FIGURE D.3: Article flow of literature search (Kidney)

Appendix E

Flowcharts of the abnormality thresholds within ACR

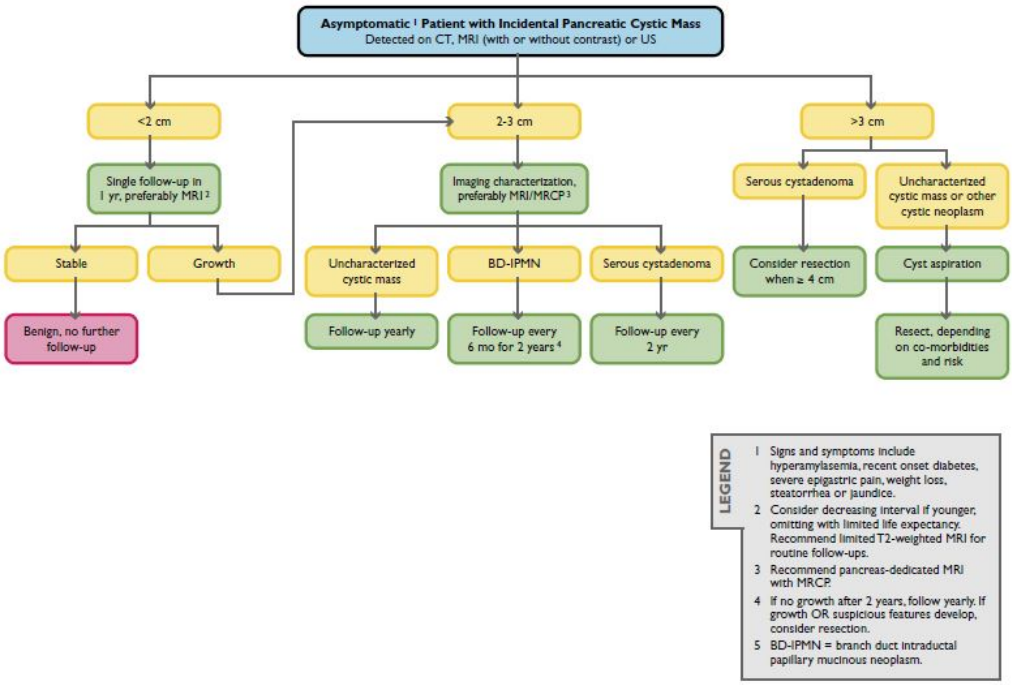


FIGURE E.1: Pancreas thresholds (ACR)

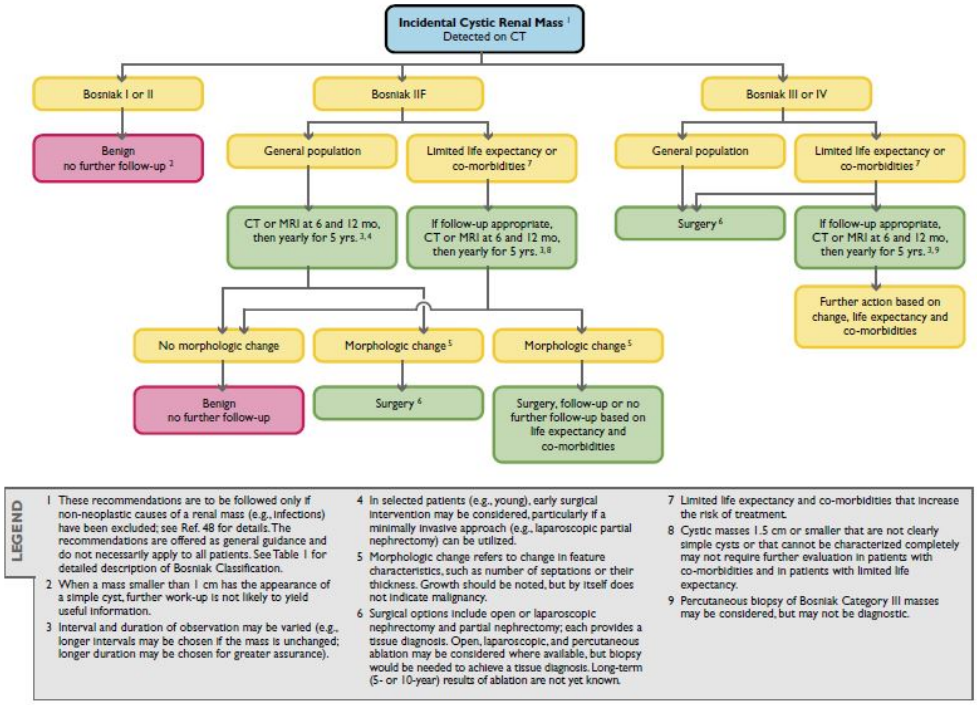


FIGURE E.2: Cystic renal masses thresholds (ACR)

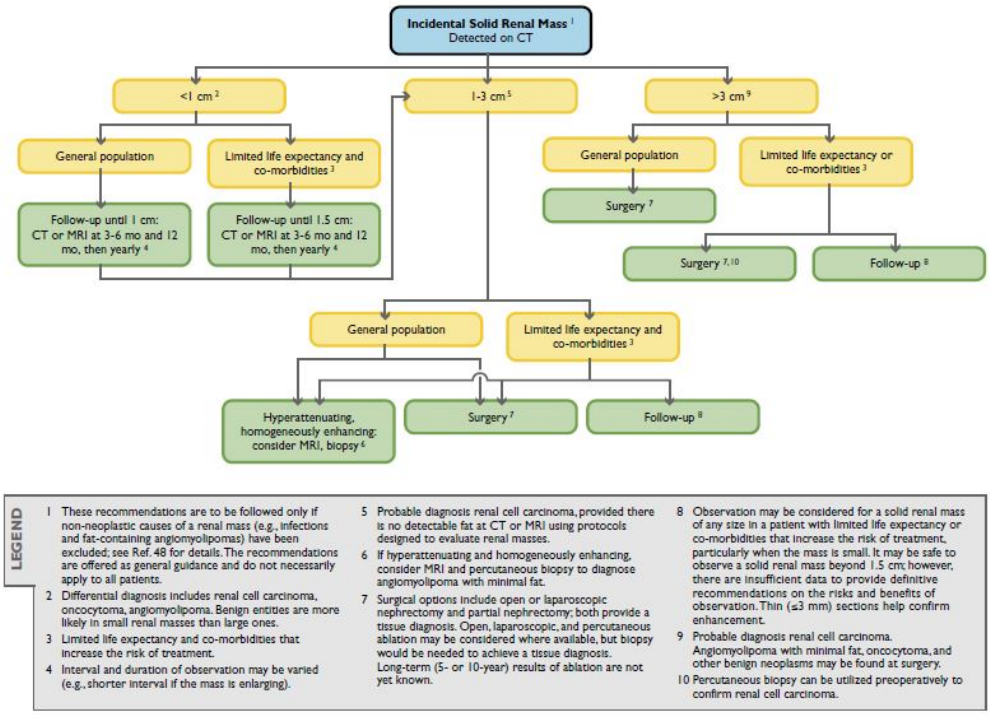


FIGURE E.3: Solid renal masses thresholds (ACR)

Appendix F

Output tables of IR calculations

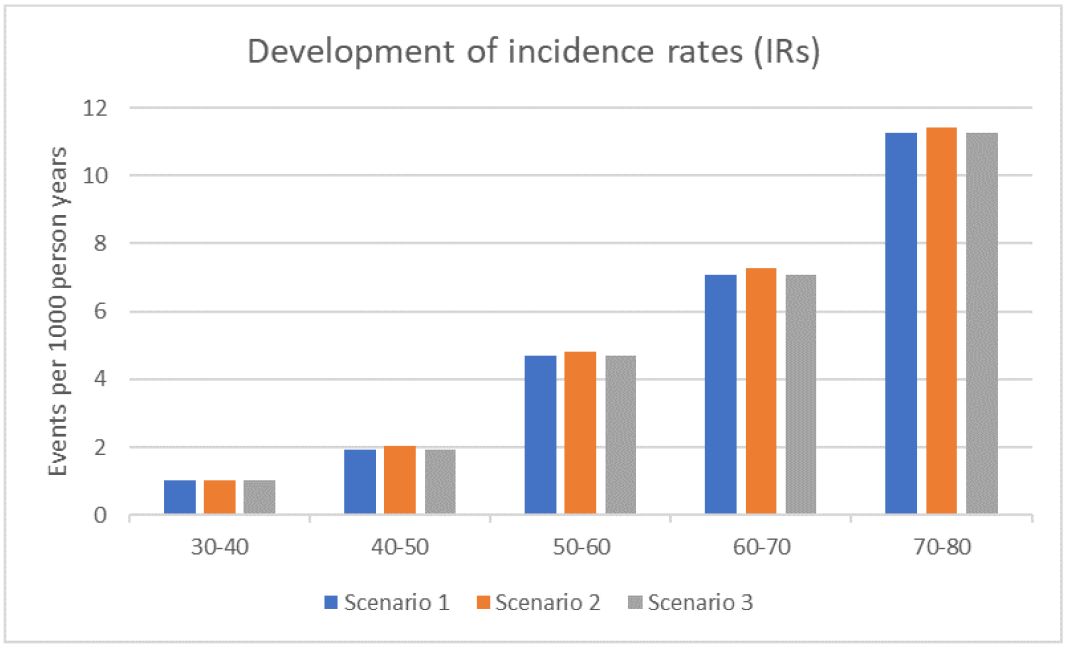


FIGURE F.1: Incidence rate development between ages 30 and 80 (bins groups of 10 years)

IR1	IR2	IR3	LB	UB	Amount of subjects (n)
0.00	0.00	0.00	20	30	[redacted]
1.02	1.02	1.02	30	40	[redacted]
1.94	2.04	1.94	40	50	[redacted]
4.69	4.81	4.69	50	60	[redacted]
7.07	7.26	7.06	60	70	[redacted]
11.27	11.41	11.26	70	80	[redacted]
24.09	24.09	24.09	80	90	[redacted]

TABLE F.1: Model output of incidence rate development, 10 years increase (ages 20 up to 100)

IR1	IR2	IR3	LB	UB	Amount of subjects (n)
1.54	1.54	1.54	30	35	[redacted]
2.11	2.11	2.11	35	40	[redacted]
1.83	2.12	1.83	40	45	[redacted]
4.43	4.74	4.45	45	50	[redacted]
6.44	6.56	6.42	50	55	[redacted]
8.27	8.65	8.24	55	60	[redacted]
10.62	11.07	10.60	60	65	[redacted]
10.58	11.20	10.55	65	70	[redacted]
11.78	11.78	11.77	70	75	[redacted]
23.45	24.08	23.27	75	80	[redacted]

TABLE F.2: Model output of incidence rate development, 5 years increase (ages 30 up to 80)

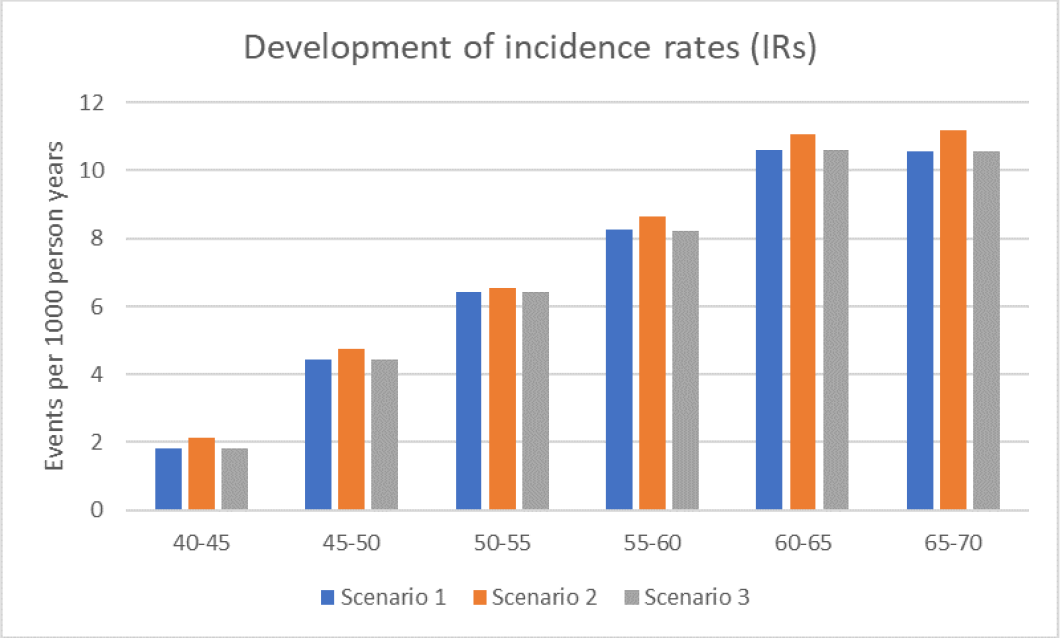


FIGURE F.2: Incidence rate development between ages 40 and 70 (bins groups of 5 years)

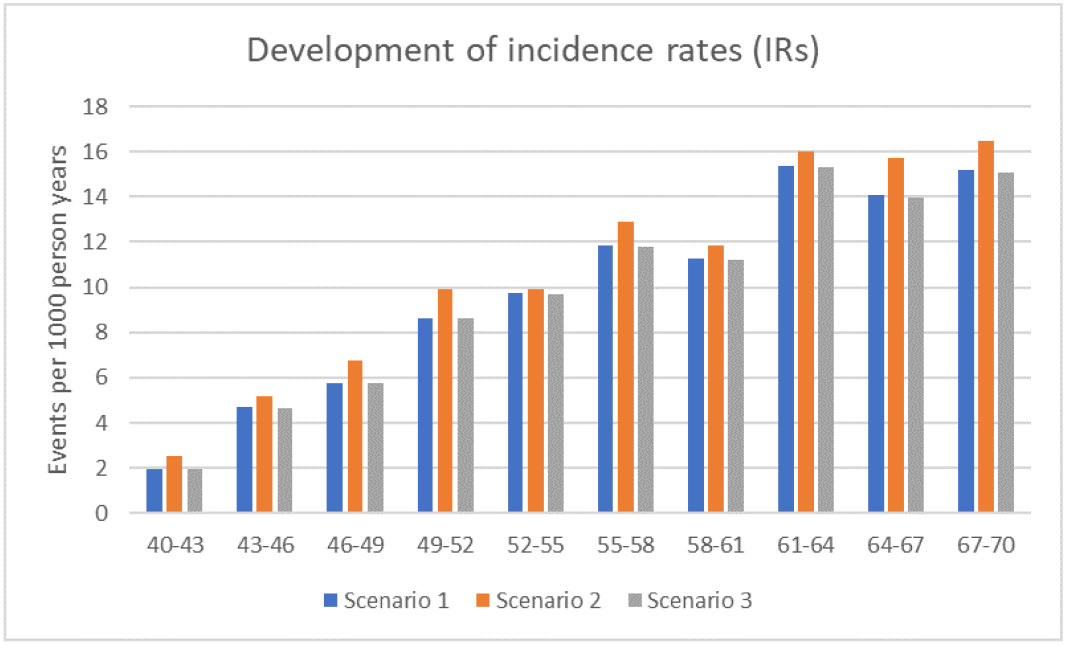


FIGURE F.3: Incidence rate development between ages 40 and 70 (bins groups of 3 years)

IR1	IR2	IR3	LB	UB	Amount of subjects (n)
28.83	32.44	28.34	50	51	[redacted]
18.21	26.18	18.10	51	52	[redacted]
27.74	32.28	27.49	52	53	[redacted]
20.73	22.03	20.33	53	54	[redacted]
30.40	30.66	30.26	54	55	[redacted]
37.27	43.17	36.20	55	56	[redacted]
30.11	37.09	29.54	56	57	[redacted]
29.27	31.26	29.26	57	58	[redacted]
21.77	30.54	21.50	58	59	[redacted]
38.87	39.29	38.17	59	60	[redacted]

IR1	IR2	IR3	LB	UB	Amount of subjects (n)
31.15	32.96	30.54	60	61	[redacted]
52.30	53.49	52.18	61	62	[redacted]
42.34	47.86	42.15	62	63	[redacted]
28.55	29.74	28.17	63	64	[redacted]
40.77	46.76	40.33	64	65	[redacted]
34.89	42.37	34.71	65	66	[redacted]
33.21	39.37	32.69	66	67	[redacted]
32.52	37.90	32.29	67	68	[redacted]
49.93	55.87	48.49	68	69	[redacted]
36.05	42.36	36.03	69	70	[redacted]

IR1	IR2	IR3	LB	UB	Amount of subjects (n)
42.15	50.17	40.37	70	71	[redacted]
25.30	25.30	25.30	71	72	[redacted]
36.99	46.24	33.62	72	73	[redacted]
35.17	32.77	32.51	73	74	[redacted]
62.37	62.37	62.37	74	75	[redacted]
66.59	66.59	66.59	75	76	[redacted]
51.34	51.34	51.34	76	77	[redacted]
126.02	160.81	123.12	77	78	[redacted]
68.86	68.86	68.86	78	79	[redacted]
29.35	29.35	29.35	79	80	[redacted]
47.12	47.12	47.12	80	81	[redacted]

TABLE F.3: Model output of incidence rate development of remaining subgroups with 1 year increase (starting ages 50 up to 70)

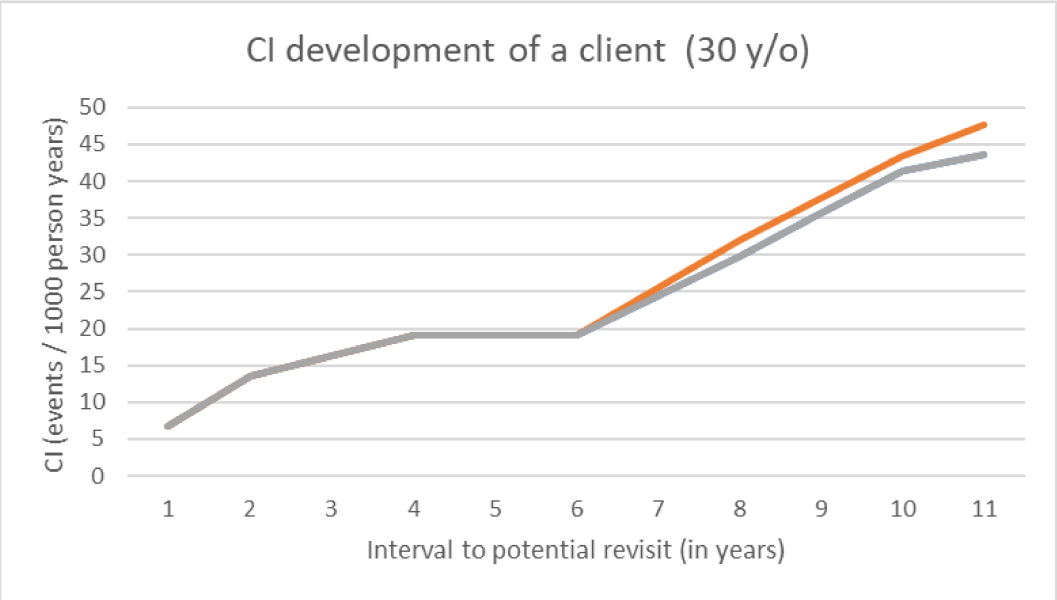


FIGURE F.4: Cumulative incidence rate and related interval period for revisit with starting age 30

IR1	IR2	IR3	LB	UB	Amount of subjects (n)
6.78	6.78	6.78	30	32	[redacted]
2.74	2.74	2.74	32	34	[redacted]
0.00	0.00	0.00	34	36	[redacted]
5.39	6.43	5.39	36	38	[redacted]
5.78	5.78	5.78	38	40	[redacted]
2.24	4.09	2.23	40	42	[redacted]

TABLE F.4: Model output of incidence rate development, 2 year increase (ages 31-41)

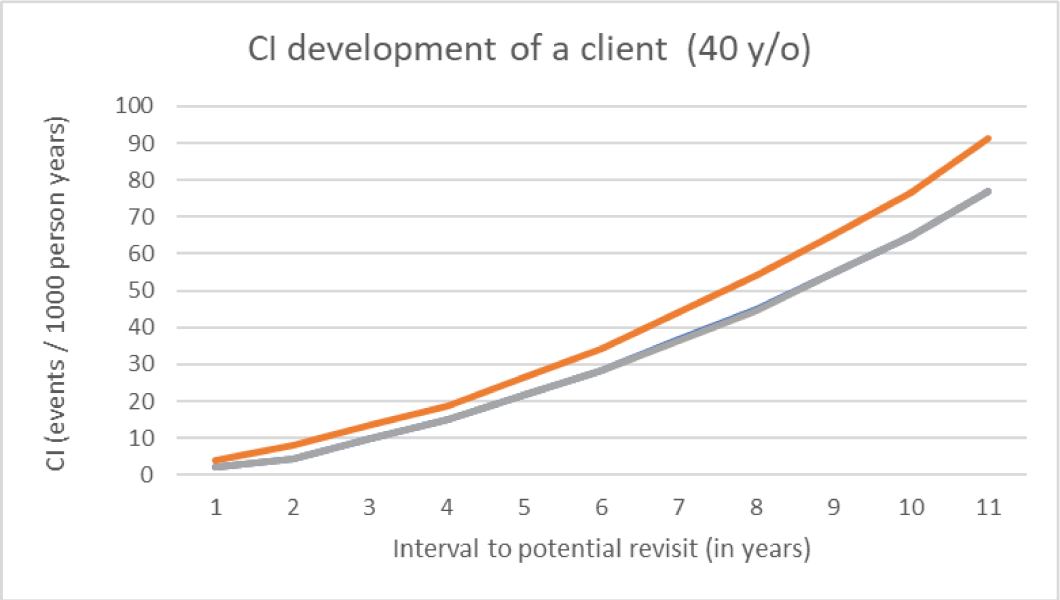


FIGURE F.5: Cumulative incidence rate and related interval period for revisit with starting age 40

IR1	IR2	IR3	LB	UB	Amount of subjects (n)
2.24	4.09	2.23	40	42	[redacted]
5.24	5.24	5.24	42	44	[redacted]
6.70	7.68	6.66	44	46	[redacted]
8.23	10.01	8.19	46	48	[redacted]
10.06	11.21	10.03	48	50	[redacted]
12.15	14.63	12.07	50	52	[redacted]

TABLE F.5: Model output of incidence rate development, 2 year increase (ages 41-51)

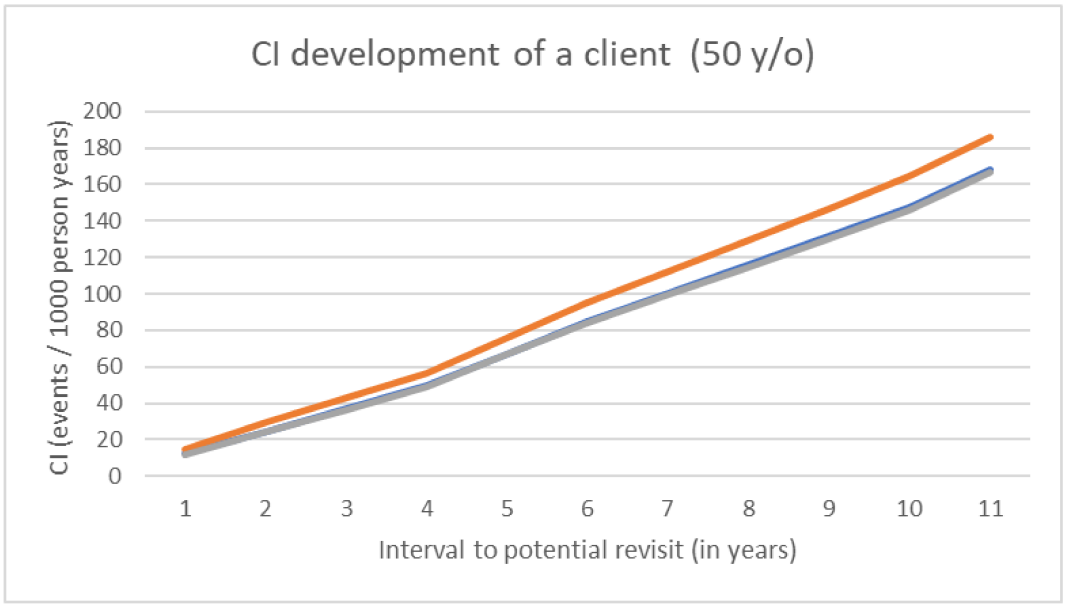


FIGURE F.6: Cumulative incidence rate and related interval period for revisit with starting age 50

IR1	IR2	IR3	LB	UB	Amount of subjects (n)
12.15	14.63	12.07	50	52	[redacted]
12.60	13.58	12.51	52	54	[redacted]
17.53	19.23	17.38	54	56	[redacted]
15.55	17.27	15.42	56	58	[redacted]
15.67	17.35	15.50	58	60	[redacted]
21.26	21.75	21.10	60	62	[redacted]

TABLE F.6: Model output of incidence rate development, 2 year increase (ages 51-61)

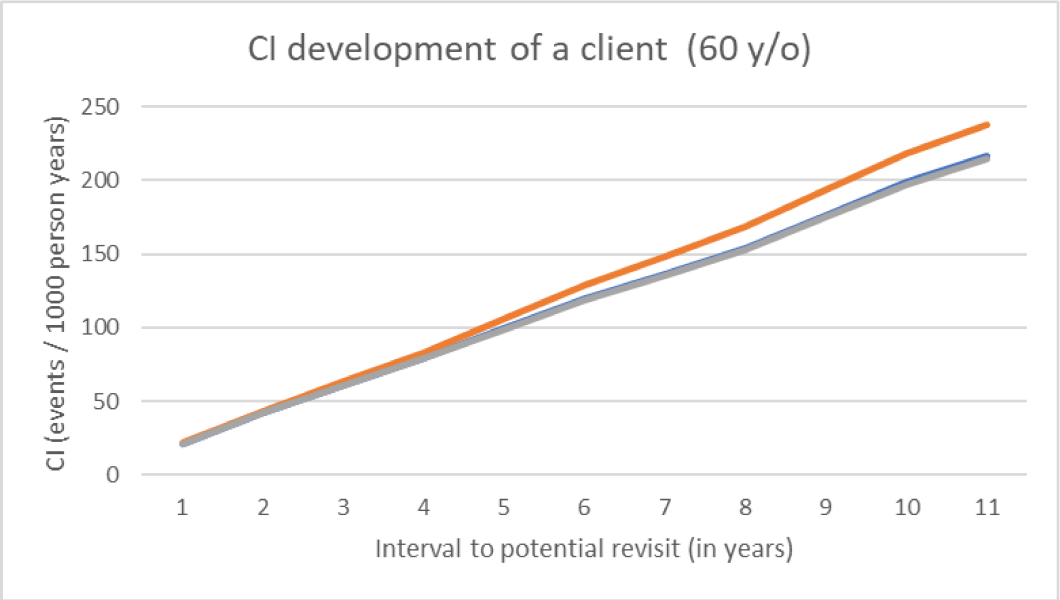


FIGURE F.7: Cumulative incidence rate and related interval period for revisit with starting age 60

IR1	IR2	IR3	LB	UB	Amount of subjects (n)
21.26	21.75	21.10	60	62	[redacted]
18.43	19.74	18.31	62	64	[redacted]
19.97	22.83	19.84	64	66	[redacted]
17.16	19.69	17.01	66	68	[redacted]
22.59	25.20	22.32	68	70	[redacted]
17.97	18.78	17.68	70	72	[redacted]

TABLE F.7: Model output of incidence rate development, 2 year increase (ages 61-71)

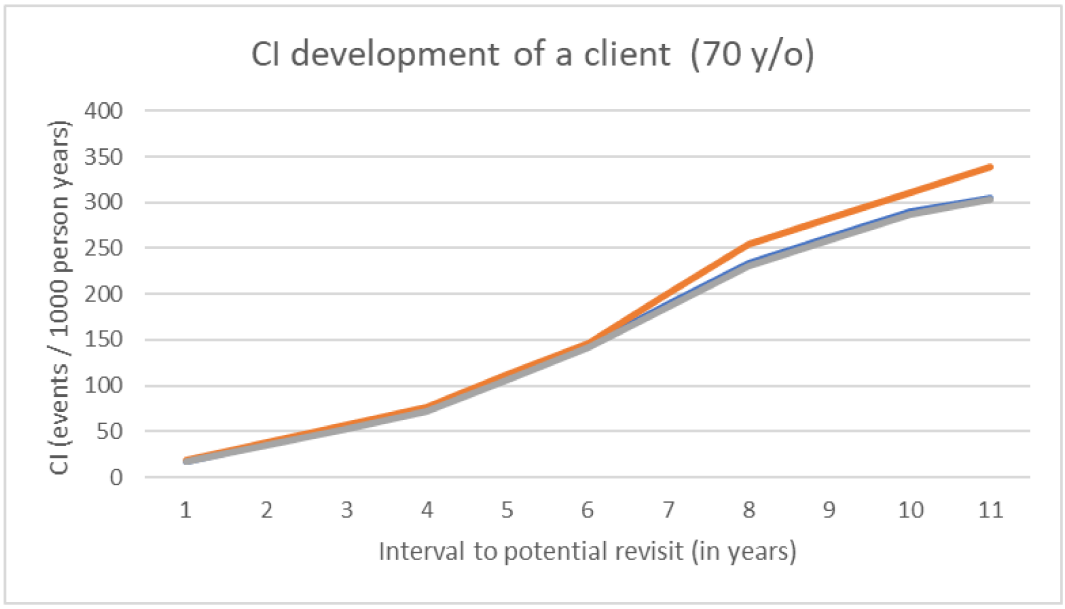


FIGURE F.8: Cumulative incidence rate and related interval period for revisit with starting age 70

IR1	IR2	IR3	LB	UB	Amount of subjects (n)
17.97	18.78	17.68	70	72	[redacted]
18.83	19.45	18.14	72	74	[redacted]
34.89	34.89	34.89	74	76	[redacted]
45.22	54.14	44.96	76	78	[redacted]
27.74	27.74	27.74	78	80	[redacted]
16.08	28.88	15.99	80	82	[redacted]

TABLE F.8: Model output of incidence rate development, 2 year increase (ages 71-81)

Appendix G

Confidential chapter

This chapter is not available in public version.