

Model-based security analysis in seven steps — a guided tour to the CORAS method

F den Braber, I Hogganvik, M S Lund, K Stølen and F Vraalsen

This paper presents the CORAS method for model-based security analysis. The presentation is case-driven. We follow two analysts in their interaction with an organisation by which they have been hired to carry out a security risk analysis. The analysis is divided into seven main steps, and the paper devotes a separate section to each of them. The paper focuses in particular on the use of the CORAS security risk modelling language as a means for communication and interaction during the seven steps.

1. Introduction

Have you ever asked yourself some of the following questions.

- Should I worry when using my credit card on the Internet?
- How safe is my Internet bank account?
- How many doctors or healthcare personnel have access to my personal health records?
- Can I be sure that I am the only one reading my e-mail?
- How crucial can a single personal mistake be for my company?

A security risk analysis may provide answers to such questions. CORAS is a method for conducting security risk analysis, which is abbreviated to 'security analysis' in the rest of this paper. CORAS provides a customised language for threat and risk modelling, and comes with detailed guidelines explaining how the language should be used to capture and model relevant information during the various stages of the security analysis. In this respect CORAS is model-based. The Unified Modelling Language (UML) [1] is typically used to model the target of the analysis. For documenting intermediate results, and for presenting the overall conclusions we use special CORAS diagrams which are inspired by UML. The CORAS method provides a computerised tool designed to support documenting, maintaining and reporting analysis results through risk modelling, table-based documentation, consistency checking and more.¹

¹ The tool may be downloaded from <http://coras.sourceforge.net/>

The seven steps of security analysis in CORAS are summarised as follows.

- Step 1

The first step involves an introductory meeting. The main item on the agenda for this meeting is to get the representatives of the client to present their overall goals of the analysis and the target they wish to have analysed. Hence, during the initial step the analysts will gather information based on the client's presentations and discussions.

- Step 2

The second step also involves a separate meeting with representatives of the client. However, this time the analysts will present their understanding of what they learned at the first meeting and from studying documentation that has been made available to them by the client. The second step also involves a rough, high-level security analysis. During this analysis the first threats, vulnerabilities, threat scenarios and unwanted incidents are identified. They will be used to help with directing and scoping the more detailed analysis still to come.

- Step 3

The third step involves a more refined description of the target to be analysed, and also all assumptions and other preconditions being made. Step three is terminated once all this documentation has been approved by the client.

- Step 4

This step is organised as a workshop, drawn from people with expertise on the target of the analysis. The goal is

to identify as many potential unwanted incidents as possible, as well as threats, vulnerabilities and threat scenarios.

- Step 5

The fifth step is also organised as a workshop. This time with the focus on estimating consequences and likelihood values for each of the identified unwanted incidents.

- Step 6

This step involves giving the client the first overall risk picture. This will typically trigger some adjustments and corrections.

- Step 7

The last step is devoted to treatment identification, as well as addressing cost/benefit issues of the treatments. This step is best organised as a workshop.

In the following sections we will give a presentation of the seven steps (see Fig 1) by means of an example where the CORAS method is applied to a telemedicine case². We distinguish the example from general descriptions of the method, language, etc, by formatting the example in italic. We end each section with a summary of the process and guidelines for the relevant step. In addition to sections for the seven steps, there is also a summary and comparison to related work. Before we start, let us introduce the example that will follow us throughout this paper.

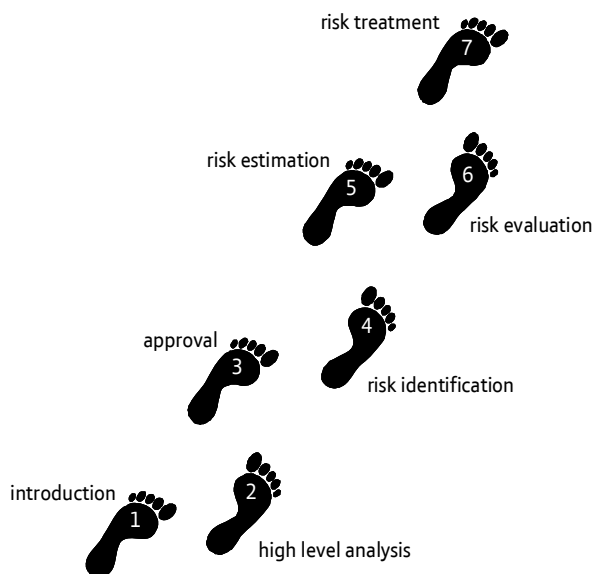


Fig 1 Steps on 'the CORAS tour'.

² This example is inspired by security analyses of real telemedicine systems conducted within the CORAS project [2, 3]. However, changes have been made so that it no longer represents any real case.

In one region of the country, an experimental telemedicine system has been set up. A dedicated network between the regional hospital and several primary healthcare centres (PHCCs) allows a general practitioner (GP) to conduct a cardiological examination of a patient (at the PHCC) in co-operation with a cardiologist located at the hospital. During an examination both medical doctors have access to the patient's health record and all data from the examination is streamed to the cardiologist's computer.

The National Ministry of Health is concerned whether the patient privacy is sufficiently protected, and hires a security analysis consultancy company to do a security analysis of the cardiology system. The consultancy company appoints a security analysis leader and a security analysis secretary to do the job (referred to as 'the analysts'). In co-operation with the ministry, they organise the meetings and workshops of which the security analysis is comprised.

2. Step 1 — introductory meeting

Before starting to identify and analyse potential risks to something, it is necessary to know exactly what this something is. What is the scope and what are the assumptions that we make? In other words, we need to know what we are supposed to protect before we can start finding what threatens it and how it may be harmed, as well as how it should be protected. It is essential that the client of the security analysis and the analysts obtain a common understanding of the terminology to be used, the target of the analysis, the assets to be protected, and the scope of the analysis.

Let us see what this means in practice for the telemedicine case.

A meeting is organised where, in addition to the analysts and a representative from the ministry, the IT manager of the regional hospital and a general practitioner from one of the PHCC's participate.

This meeting is where the overall setting of the analysis is decided, and the first step is taken towards establishing the target description that will be used later in the analysis. The meeting starts with the security analysis leader giving a brief presentation of the method to be used, what the client (the National Ministry of Health) can expect from the analysis, and a proposed meeting plan. The analysis leader reminds the client of the responsibilities with respect to providing necessary information and documentation about the target in question, as well as allocating people with suitable backgrounds to participate in the security analysis meetings and workshops.

The IT manager then presents the telemedicine system which will be the target of analysis. As part of the presentation she draws the picture shown in Fig 2. From

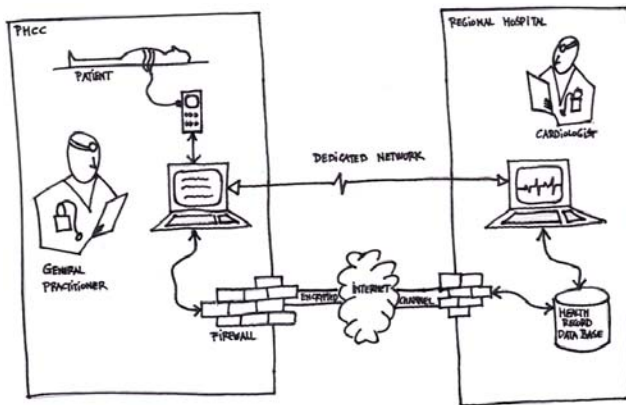


Fig 2 Picture of the target.

the picture we see that speech and other data from the examination of a patient is streamed over a dedicated network, while access to the patient's health record (stored in a database at the regional hospital) is given through an encrypted channel over the Internet. Next in line after the IT manager is the medical doctor from the PHCC. She talks about her personal experiences from using the system.

After the presentations, a discussion on the scope and focus of the analysis follows. The representative of the ministry emphasises that they are particularly worried about the confidentiality and integrity of the health records and other medical data, first and foremost for the sake of the patients' health, but also because of the public's trust in the national healthcare system. For the medical doctor the most important thing is the patient's health and well-being, and hence the availability and integrity of the telemedicine system. The IT manager explains that they have already made a security analysis of the health record database and the encrypted access, so she is confident that this part of the system is secure and reliable. After some discussion the representative of the ministry decides that the focus will be on confidentiality and integrity of medical data, and the availability of the service, but that the access to the health record database is outside the scope of analysis.

As the last point on the agenda, the participants set up a plan for the rest of the analysis with dates and indications of who should be present.

Step 1 — summary

Tasks:

- the security analysis method is introduced,
- the client presents the goals and the target of the analysis,
- the focus and scope of the analysis is set,
- the meetings and workshops are planned.

People that should participate:

- analysis leader (required),
- analysis secretary (required),
- representatives of the client:
 - decision makers (required),
 - technical expertise (optional),
 - users (optional).

Modelling guideline:

- system description:
 - at this stage of the analysis it can be useful to describe the target with informal models like drawings, pictures or sketches on a blackboard,
 - the presentation can later be supplemented with more formal modelling techniques such as UML or data-flow diagrams.

3. Step 2 — high-level analysis

The second step is called the high-level analysis, and as the name indicates this involves conducting an initial analysis of the target. This step also typically involves a meeting between the analysts and the representatives of the client. The main purpose is to identify assets and get an overview of the main risks. Finding the assets that need protection is initiated in step 2 and completed in step 3. The remaining four steps of the analysis will be directed towards these assets. The outcome of the high-level analysis helps the analysts to identify the aspects of the target having the most urgent need for in-depth analysis, and hence makes it easier to define the exact scope and focus of the full analysis.

The second meeting starts with the security analysis leader presenting the analysts' understanding of the target to be analysed. The information presented by the client at the previous meeting, as well as documentation received in the mean time, has been formalised in UML diagrams [1]. The UML class diagram (Fig 3) shows the relevant concepts and how they relate, while the UML collaboration diagram (Fig 4) illustrates the physical organisation of the target. Furthermore, the medical doctor's description of use has been captured as a UML activity diagram (Fig 5). During this presentation the participants representing the client make corrections and eliminate errors, so that the result is a target description all parties can agree upon. In the class and collaboration diagrams the security analysis leader has also indicated what areas are understood to be the focus of the analysis.

After agreeing on a target description, the analysis moves on to asset identification. An asset is something in or related to the target to which the client assigns great value. Based on the discussion at the introductory meeting, the analysis leader has prepared an initial 'CORAS asset diagram' (Fig 6) to help with specifying the scope of the analysis. The asset diagram shows the National Ministry of Health as the client (i.e. the

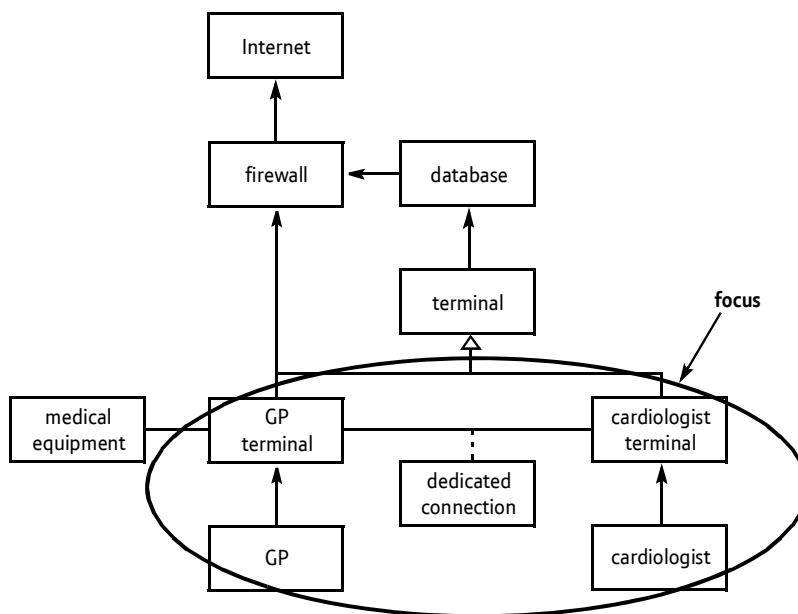


Fig 3 Class diagram showing a conceptual view of the target.

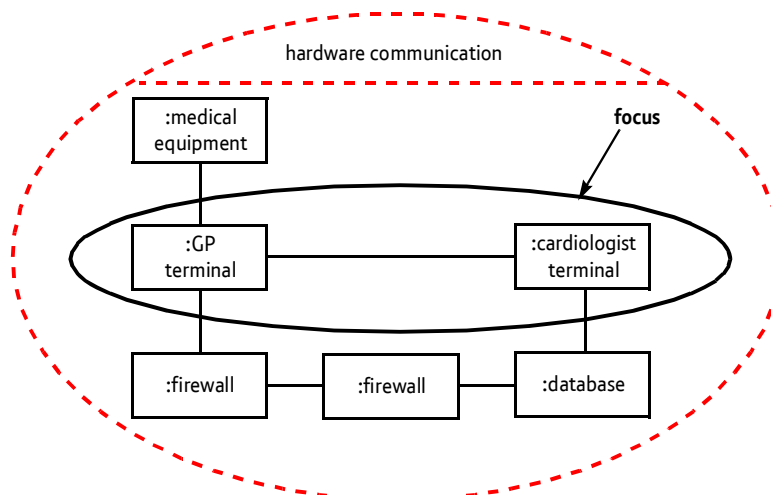


Fig 4 Collaboration diagram illustrating the physical communication lines.

stakeholder that is initiating and paying for the analysis), and its four assets: ‘Health records’, ‘Provision of telecardiology service’, ‘Patient’s health’ and ‘Public’s trust in system’. Because trust and health are difficult to measure, especially in a technical setting like this, the analysis leader makes a distinction between direct and indirect assets. He explains direct assets as assets that may be harmed directly by an unwanted incident, while the indirect assets are only harmed if one of the direct assets is harmed first. In the asset diagram the direct assets are placed within the target of analysis region and the indirect are placed outside.

The arrows show dependencies between the assets, such that, for example harm to ‘Health records’ may cause harm to ‘Public’s trust in system’. The dashed lines

in Fig 6 symbolise the client’s, or other interested parties’, relation to the assets.

After agreeing on the assets, the analysts conduct a high-level analysis together with the analysis participants. The short brainstorming should identify the most important threats and vulnerabilities, but without going into great detail. In this case the client is concerned about hackers, eavesdroppers, system failure and whether the security mechanisms are sufficient.

These threats and vulnerabilities do not necessarily involve major risks, but give the analysis leader valuable input on where to start the analysis. The analysis secretary documents the results by filling in the high-level risk table shown in Table 1.

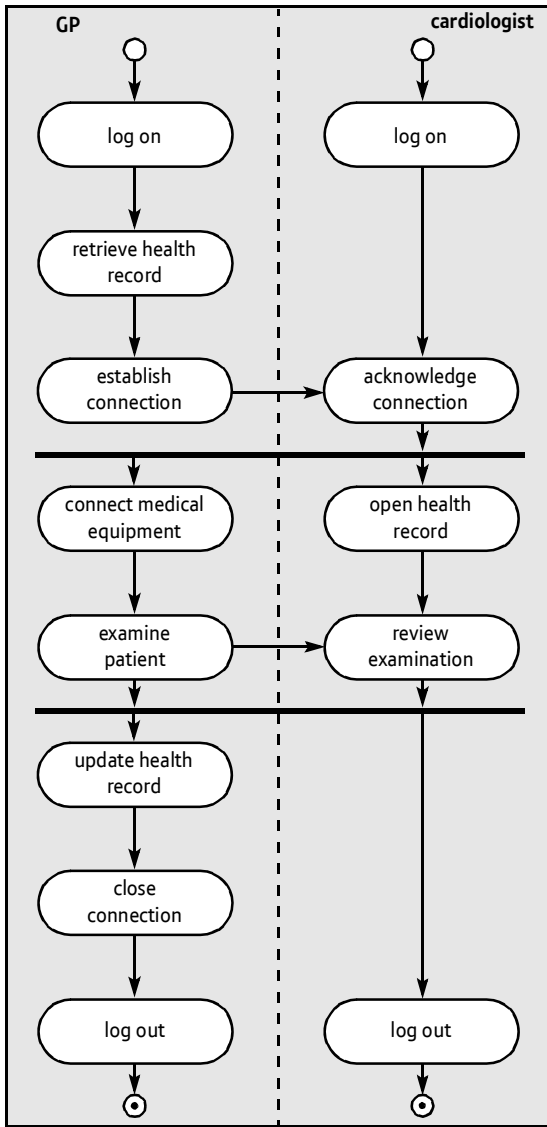


Fig 5 Activity diagram describing the parallel processes of the GP and the cardiologist.

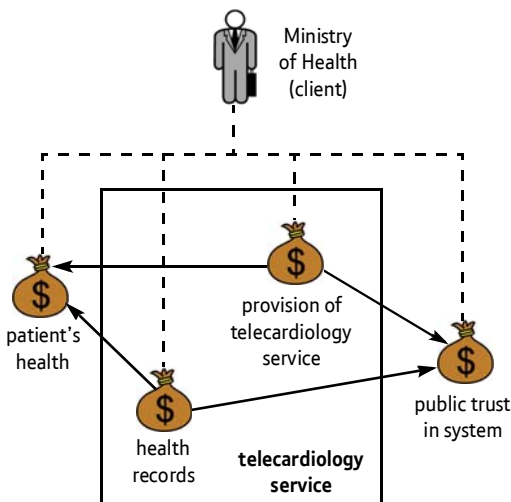


Fig 6 Asset diagram.

Step 2 — summary

Tasks:

- the target as understood by the analysts is presented,
- the assets are identified,
- a high-level analysis is conducted.

People that should be present:

- security analysis leader (required),
- security analysis secretary (required),
- representatives of the client:
 - decision makers (required),
 - technical expertise (required),
 - users (optional).

Modelling guidelines:

- asset diagrams:
 - draw a region that logically or physically represents the target of analysis,
 - place the direct assets within the region,
 - place the indirect assets outside the region (indirect assets are harmed as a consequence of a direct asset being harmed first),
 - indicate with arrows which assets may affect other assets,
 - assets may be ranked according to their importance,
 - if the analysis has more than one client, the clients should be associated with their assets,
- target descriptions:
 - use a formal or standardised notation such as UML [1], but ensure that the notation is explained thoroughly so that the participants understand it,
 - create models of both the static and the dynamic features of the target (static may be hardware configurations, network design, etc, while dynamic may be work processes, information flow, etc),
 - for the static parts of the description UML class diagrams and UML collaboration diagrams (or similar notations) are recommended,
 - for the dynamic parts we recommend UML activity diagrams and UML sequence diagrams (or similar notations).

Table 1 High-level risk table.

	<p>threat (accidental)</p> <p>threat (deliberate)</p> <p>threat (non-human)</p> <p>threat scenario</p> <p>unwanted incident</p> <p>asset</p> <p>vulnerability</p>	
Who/what causes it?	How? What is the incident? What does it harm?	What makes it possible?
Hacker	Breaks into the system and steals health records	Insufficient security
Employee	Sloppiness compromises confidentiality of health records	Insufficient training
Eavesdropper	Eavesdropping on dedicated connection	Insufficient protection of connection
System failure	System goes down during examination	Unstable connection/immature technology
Employee	Sloppiness compromises integrity of health record	Prose-based health records (i.e. natural language)
Network failure	Transmission problems compromise integrity of medical data	Unstable connection/immature technology
Employee	Health records leak out by accident — compromises their confidentiality and damages the trust in the system	Possibility of irregular handling of health records

4. Step 3 — approval

The last of the preparatory steps is the approval step. The approval is often conducted as a separate meeting, but may also take place via e-mail. The main goal is to finalise the documentation and characterisation of target and assets, and get this formally approved by the client. At the end of this meeting there should be a document (possibly with a list of required changes) to which all parties agree and commit. The approval also involves defining consequence scales (for each asset) and a likelihood scale. Multiple consequence scales are used when it is difficult or inappropriate to measure damage to all assets according to the same scale, e.g. it is easier to measure ‘income’ in monetary values than ‘company brand’.

There should only be one likelihood scale appropriate for the analysis scope, e.g. based on a time-interval (years, weeks, hours, etc) or probabilities. The last activity of the approval is to decide upon the risk evaluation criteria. The criteria states which level of risk the client accepts for each of the assets.

The security analysis leader has updated the presentation from the last meeting based on comments from the other participants, and the target and asset descriptions are now approved. Based on the discussions in the first two meetings and issues identified in the high-level analysis, it is decided to narrow the scope of the analysis, and agree upon the following target definition.

The target of analysis will be the availability of the telecardiology service, and confidentiality and integrity of health records and medical data in relation to use of the service and related equipment. The indirect asset ‘Public’s trust in system’ is to be kept outside the scope.

A risk is the potential for an unwanted incident to have an impact upon objectives (assets) [4], or in other words to reduce the value of at least one of the identified assets. Often the client accepts some risks that are not judged to be critical rather than eliminating or reducing them. This may be because of shortage of resources to implement changes, conflicting concerns, or the treatment costs will be greater than the benefits. As a first step towards distinguishing risks that can be accepted from those that cannot, the representatives from the client are asked to rank the assets according to their importance (1 = very important, 5 = minor importance) and fill in the asset table (Table 2). Then the final treatment step can address the risks for the most important asset first.

Having finished the asset table, they go on to define the likelihood scale (a general description of frequency or probability [4]) of which incidents occur, and the impact or consequence they have on the assets. The analysts initiate the discussion by suggesting a scale of likelihood based on the following rule of thumb — the lower incident likelihood ‘rare’ is set to be a maximum of one occurrence during the target’s lifetime; the remaining

Table 2 Asset table.

Asset	Importance	Type
Health records	2	Direct asset
Provision of telecardiology service	3	Direct asset
Public’s trust in system	(Scoped out)	Indirect asset
Patient’s health	1	Indirect asset

intervals have an increasing number of expected events until the maximum possible number of incidents per year is reached. Because assets of different types are involved, they make separate consequence scales for each of the direct assets. Table 3 shows the consequence scale defined for the asset 'Health records' in terms of number of health records affected. If feasible, the consequence description for an asset may include more than one measure, e.g. 'major' could be the number of disclosed health records, or the number of deleted records, etc. Table 4 gives the likelihood scale defined for the target as such. By using the same scale for all scenarios and incidents, it is possible to extract combined likelihood values as shown later in the risk estimation step.

Table 3 Consequence scale for 'health records'.

Consequence value	Description
Catastrophic	1000+ health records (HRs) are affected
Major	100-1000 HRs are affected
Moderate	10-100 HRs are affected
Minor	1-10 HRs are affected
Insignificant	No HR is affected

Table 4 Likelihood scale.

Likelihood value	Description ³
Certain	Five times or more per year (50-*: 10y = 5-*: 1y)
Likely	Two to five times per year (21-49: 10y = 2,1-4,9: 1y)
Possible	Once a year (6-20: 10y = 0,6-2: 1y)
Unlikely	Less than once per year (2-5: 10y = 0,2-0,5: 1y)
Rare	Less than once per ten years (0-1:10y = 0-0,1:1y)

Finally, the representatives of the client need to define the risk evaluation criteria, the criteria which assert whether a risk to an asset is acceptable or whether it is necessary to evaluate possible treatments for it. They define these criteria by means of a risk evaluation matrix for each asset. The security analysis leader draws the matrix for the asset 'Health records' on a blackboard. It has likelihood and consequence values as its axes so that a risk with a specific likelihood and consequence will belong to the intersecting cell. Based on a discussion in the group, the security analysis leader marks the cells in the matrix as 'acceptable' or 'must be evaluated'. The resulting risk evaluation matrix is shown in Table 5, and

³ 50-*:10y is short for 50 or more incidents per 10 years, equivalent to 5 or more incidents per year.

Table 5 Risk evaluation matrix.

		Consequence				
		Insignificant	Minor	Moderate	Major	Catastrophic
Frequency	Rare	Acceptable	Acceptable	Acceptable	Acceptable	Must be evaluated
	Unlikely	Acceptable	Acceptable	Acceptable	Must be evaluated	Must be evaluated
	Possible	Acceptable	Acceptable	Must be evaluated	Must be evaluated	Must be evaluated
	Likely	Acceptable	Must be evaluated	Must be evaluated	Must be evaluated	Must be evaluated
	Certain	Must be evaluated	Must be evaluated	Must be evaluated	Must be evaluated	Must be evaluated

the participants decide to let this matrix cover the other assets as well.

After completing this task for all assets the analysts and the participants have the framework and vocabulary they need to start identifying threats (a potential cause of an unwanted incident [5]), vulnerabilities (weaknesses which can be exploited by one or more threats [5]), unwanted incidents and risks, and can move on to the next step.

Step 3 — summary

Tasks:

- the client approves target descriptions and asset descriptions,
- the assets should be ranked according to importance,
- consequence scales must be set for each asset within the scope of the analysis,
- a likelihood scale must be defined,
- the client must decide risk evaluation criteria for each asset within the scope of the analysis.

Participants:

- the same as in the previous meeting, but, since this step sets the boundaries for the further analysis, it is important that the relevant decision-makers are present.

5. Step 4 — risk identification

To identify risks CORAS makes use of a technique called structured brainstorming. Structured brainstorming may be understood as a structured 'walk-through' of the target of analysis and is carried out as a workshop. The main idea of structured brainstorming is that since the analysis participants represent different competences, backgrounds and interests, they will view the target from different perspectives and consequently identify more, and possibly other, risks than individuals or a more homogeneous group would have managed.

The findings from the brainstorming are documented with the CORAS security risk modelling language. We will now exemplify how we model risks with the CORAS language, using the symbols presented in Fig 7.



Fig 7 Symbols from the CORAS risk modelling language.

The analysis leader challenges the participants to work with questions such as: What are you most worried about with respect to your assets (threat scenarios and unwanted incidents)? Who/what may initiate these (threats)? What makes this possible (vulnerabilities)? This information is modelled by the secretary in threat diagrams.

The analysis leader has used this technique on numerous occasions before, but does not use exactly the same procedure in every case, adapting it to fit the target domain. Often it is useful to include checklists and 'best practices' for a specific technology or domain. In this case IT experts and medical personnel (general practitioners) must participate in the brainstorming, but some will only participate when their competences are needed for specific scenarios. Since people may be involved at different stages of the analysis, it is essential that information gathered during this session is documented in a simple and comprehensive way.

The analysis leader uses the target models from Step 2 (Figs 2, 3, 4 and 5) as input to the brainstorming session. The models are assessed in a stepwise and

structured manner and the identified unwanted incidents are documented on-the-fly (using the guidelines presented in the summary).

A set of initial threat scenario diagrams (Figs 8, 9 and 10) has been prepared by the analysis secretary on the basis of the high-level analysis table (Table 1). These represent a starting point for discussion and are often underspecified. She has decided to structure the three diagrams according to the ISO categorisation [5], describing the different types of threats — human accidental, human deliberate and non-human threats (environmental).

The threat diagram in Fig 8 shows how a combination of insufficient training or prose-based health records, and sloppiness may compromise the integrity and confidentiality of the patient's health records. The system also allows for irregular handling of health records where an employee may accidentally cause a leakage of records. A confidentiality or integrity breach may harm the health record in the sense that it is no longer secret nor correct. In the outmost consequence a faulty health record may affect the patient's health.

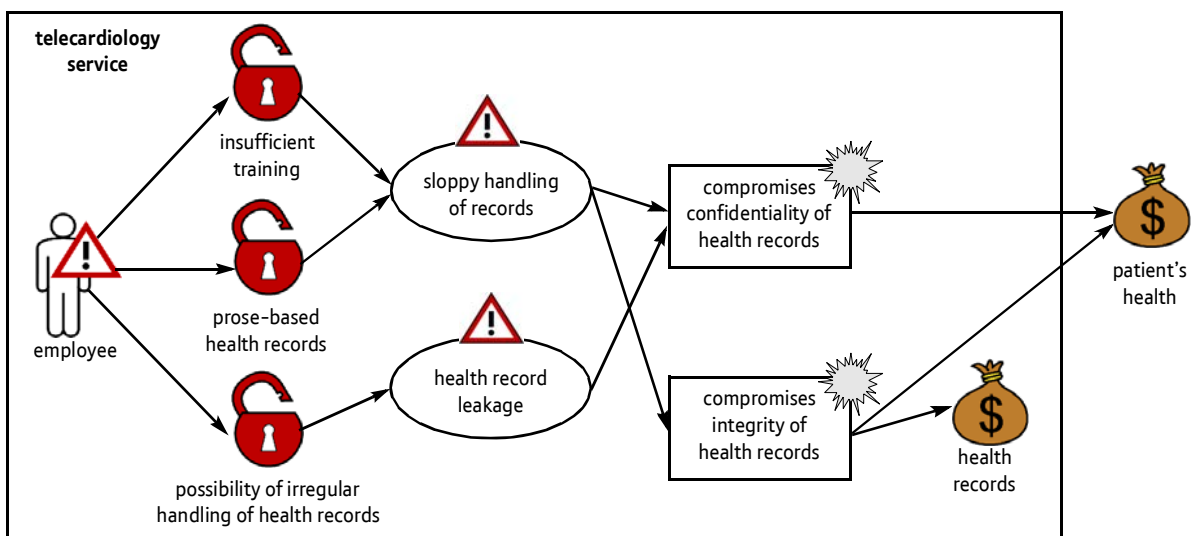


Fig 8 Initial threat diagram — accidental actions.

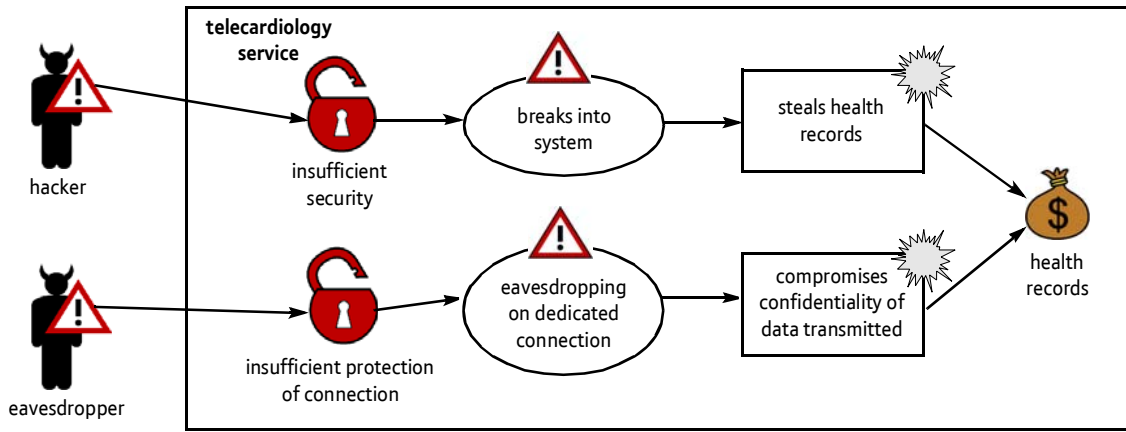


Fig 9 Initial threat diagram — deliberate actions.

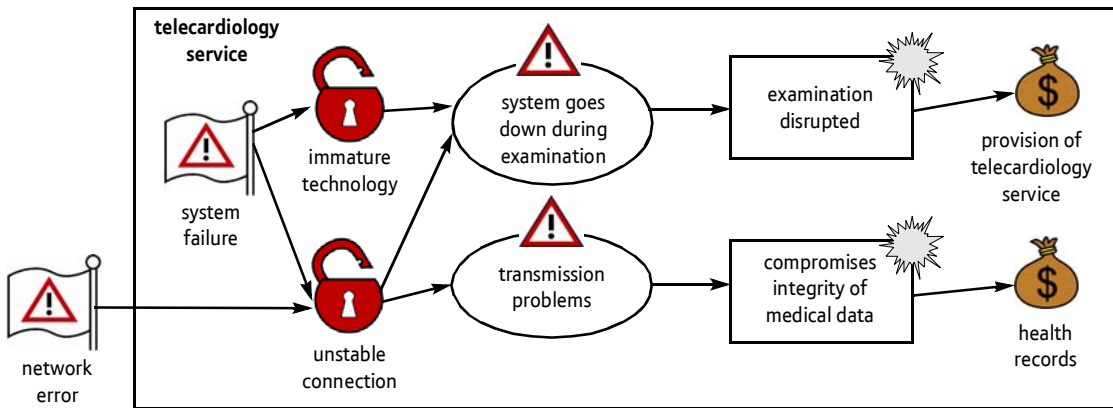


Fig 10 Initial threat diagram — non-human threats.

In the threat diagram describing deliberate harmful actions caused by humans, the participants have identified two main threats — hacker and eavesdropper (Fig 9). A hacker may exploit insufficient security mechanisms to break into the system and steal health records. An eavesdropper is a person that, due to insufficient protection of communication lines, may gather data that is transmitted and thereby compromise its confidentiality.

The participants also worry about threats like system failure and network failure (Fig 10). They fear that unstable connections or immature technology are vulnerabilities that may lead to system crashes during examination or transmission problems. A transmission problem may interfere with the data that is stored in the system and leave the health records only partly correct.

During the brainstorming session the initial threat diagrams are expanded with new information on-the-fly. If the amount of information is too large, the secretary may choose to write it down or use audiovisual equipment to make sure that nothing is missed. The diagrams may then be updated and completed after the session. The threat diagram illustrating incidents caused

by employees' accidental actions (Fig 8) receives much attention among the participants and develops into Fig 11.

Due to space limitations, we will not explore the other two threat diagrams further, but concentrate on just this one.

The participants decide that the threat 'employee' must be specified into 'general practitioner (GP)' and 'IT personnel' since they may cause different incidents. If the GP has too little security training, she may store copies of health records on a local computer. This may compromise the integrity of the records and in the worst case lead to an erroneous diagnosis of a patient. The same incidents may also occur if the GP enters wrong information in the patient's health record. The system allows for irregular handling of health records which makes it possible to accidentally send records to unauthorised people. This would compromise the confidentiality of the health record. The policy of the IT personnel with respect to access control has been very 'loose'. They explain this with their responsibility for making critical updates in emergencies and that they do not have the time to wait for a person with correct access rights to show up. An unfortunate consequence of this is that sometimes

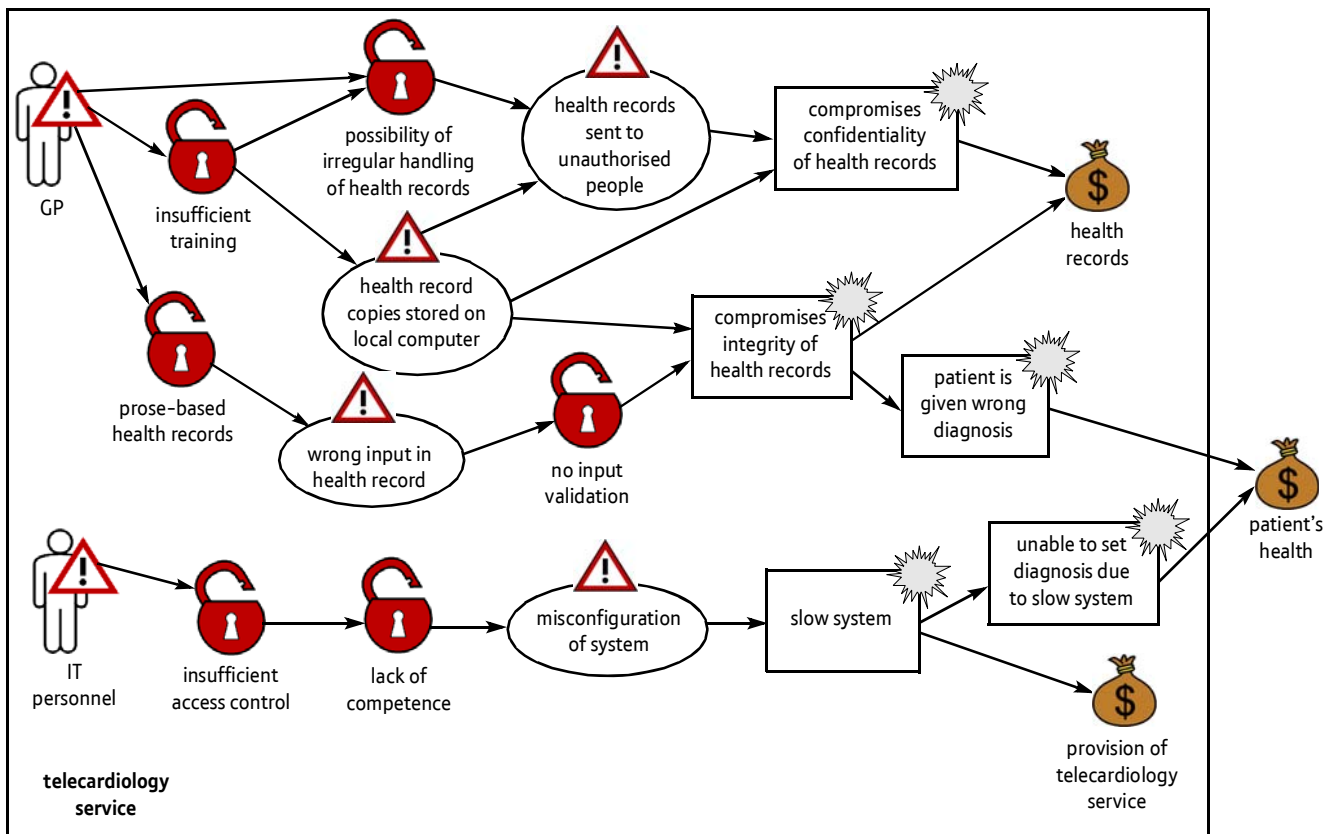


Fig 11 Final threat diagram — accidental actions.

people without the required competence become responsible for critical changes. This may lead to misconfiguration of the system, which again may slow it down. A slow system may make it impossible to set a patient's diagnosis, and also the ability to provide a telecardiology service.

Step 4 — summary

Tasks:

- the initial threat diagrams should be completed with identified threats, vulnerabilities, threat scenarios and unwanted incidents.

People that should participate:

- security analysis leader (required),
- security analysis secretary (required),
- representatives of the client:
 - decision makers (optional — because this workshop often has a technical focus and the decision makers' competence is more relevant in the next step),
 - technical expertise (required),
 - users (required) .

Modelling guideline:

- threat diagrams:

— use the region from the asset diagram and add more regions if necessary,

— model different kinds of threats in separate diagrams, e.g. deliberate sabotage in one diagram, mistakes in an other, environmental in a third, etc (the ISO/IEC standard [5] contains a useful classification) — this makes it easier to generalise the risks, e.g. 'these risks are caused by deliberate intruders' or 'these risks are caused by human errors',

— threats are placed to the left in the region, while threats that can be classified as external (hackers, intruders, etc) are placed outside the region,

— assets are listed to the right, outside the region,

— unwanted incidents are placed within the region in relation to the assets on which they have an impact,

— assets that are not harmed by any incidents are removed from the diagram,

— add threat scenarios between the threats and the unwanted incidents in the same order as they occur in real time (i.e. in a logical sequence),

— insert the vulnerabilities before the threat scenario or unwanted incident to which they lead, e.g. a vulnerability called ‘poor back-up solution’ is typically placed before the threat scenario ‘the back-up solution fails to run the application database correctly’.

6. Step 5 — risk estimation

When the threat scenarios, unwanted incidents, threats and vulnerabilities are properly described in threat diagrams it is time to estimate likelihood values and consequences. This is typically done in a separate workshop. The values are used to compute the risk value which decides whether the risk should be accepted or evaluated for treatments. The participants in the workshop provide likelihood estimates for each threat scenario in the threat diagrams. For scenarios that are difficult to estimate, the analysis leader gives suggestions based on historical data like security incident statistics or personal experience. The likelihood of the threat scenarios are used to extract a combined likelihood for unwanted incidents. Consequences are estimated for each ‘unwanted incident — asset’ relation. The consequence value is taken from the consequence scale of the asset decided in Step 3. In this workshop it is especially important to include people with the competence needed to estimate realistic likelihoods and consequences, meaning that technical expertise, users and decision makers must be included.

The analysis leader organises the estimation as a separate workshop where the input is the threat diagrams from the previous workshop. In this workshop it is especially important to include users, technical experts and decision makers to obtain estimates that are as correct as possible. The analysis participants decide that ‘most likely’ estimates will provide more realistic risk values than ‘worst case’ estimates. Firstly, they provide as many estimates as possible for the threat scenarios which help estimating the likelihood of the unwanted incidents (if this cannot be established by other means). Secondly, the consequences of the unwanted incidents for each harmed asset are estimated. The estimates are documented by annotating the diagrams as shown in Fig 12 — further details can be specified in a table.

There are different ways of computing the likelihood of an incident that may be caused by more than one threat scenario. If the estimates are suitable for mathematical calculations a computerised tool may be used. Since the likelihood scale in our case is in the form of intervals, the analysis leader decides to use an informal method that is quite straightforward and transparent. The threat scenario ‘Health records sent out to unauthorised people’ and ‘Health record copies stored on local computer’ can both lead to ‘Compromises confidentiality of health records’. Table 6 shows how the combined likelihood is estimated. The technique is informal, but suitable for the creative structured

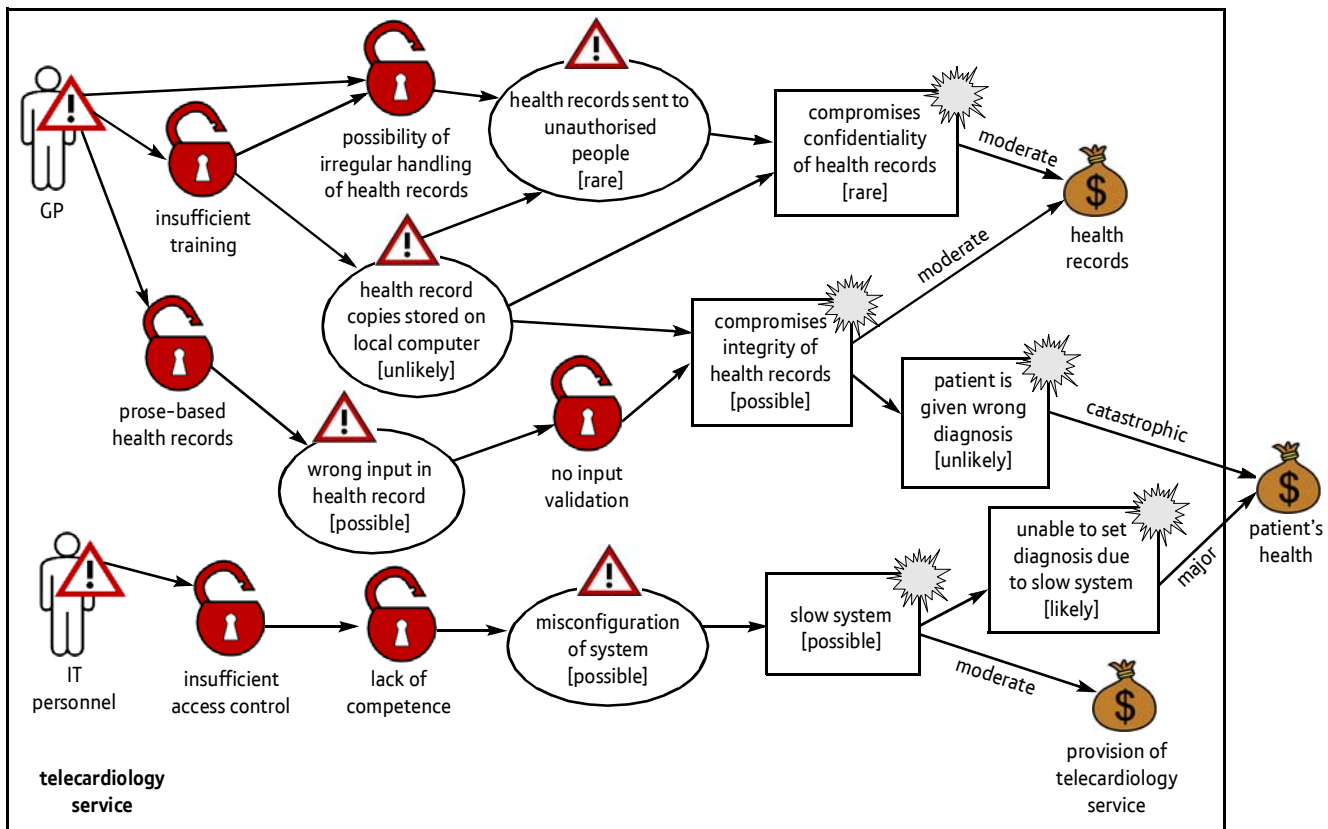


Fig 12 Threat diagram with likelihood and consequence estimates.

Table 6 Combined likelihood estimates.

Threat scenario	Likelihood	Unwanted incident	Combined likelihood
Health records sent out to unauthorised people	Rare (0-1:10y)	Compromises confidentiality of health records	(0-1:10y)+(2-5:10y)=(2-6:10y) Some overlap between unlikely and possible, but it fits best in the unlikely interval.
Health record copies stored on local computer	Unlikely (2-5:10y)		

brainstorming setting. For more precise calculation of probabilities fault tree analysis (FTA)[6] may be used. It is of course important that the combined estimates reflect reality, meaning that the combined estimates should be presented to the participants for validation.

In this case, the participants reject the suggested estimate for 'Compromises confidentiality of health records', arguing that the likelihood is less than 'unlikely' and adjust it to 'rare'.

Step 5 — summary

Tasks:

- every threat scenario must be given a likelihood estimate and unwanted incident likelihoods are based on these,
- every relation between an unwanted incident and an asset must be given a consequence estimate.

People that should be present:

- security analysis leader (required),
- security analysis secretary (required),
- representatives of the client:
 - decision makers (required),
 - technical expertise regarding the target (required),
 - users (required).

Modelling guideline:

- risk estimation on threat diagrams:
 - add likelihood estimates to the threat scenarios,
 - add likelihood estimates to the unwanted incidents, based on the threat scenarios,
 - annotate each unwanted incident-asset relation with a consequence taken from the respective asset's consequence scale.

7. Step 6 — risk evaluation

The risk evaluation consists of two activities. Firstly, the analysis secretary uses the likelihood and consequence estimates to compute the risk values and to place the risks in the risk matrix. Secondly, the resulting risk matrices are presented to the client for inspection. This presentation may be given in a separate meeting or included in the treatment workshop (Step 7).

In our case the risk value is determined by the risk evaluation matrix. From the four unwanted incidents in the threat diagram, the analysis secretary extracts five risks. 'Compromising the confidentiality of health records' (CC1) may affect health records. 'Compromising the integrity of health records' may also harm health records (CI1), in addition to patient's health if it contributes to a faulty diagnosis (PR1). Finally, 'slow system' may slow down an examination (SS2) and harm the patient's health (SS1). Only CC1 is within acceptable risk levels, the rest need further evaluation. Table 7 shows the risks placed in the risk evaluation matrix.

Table 7 Risk evaluation matrix with risks consequence.

		Consequence				
		Insignificant	Minor	Moderate	Major	Catastrophic
Likelihood	Rare			CC1		
	Unlikely					PR1
	Possible			CI1, SS2		
	Likely				SS1	
	Certain					

The analysis leader gives the participants an opportunity to adjust likelihood and consequence estimates, and risk acceptance levels, to make sure that the results reflect reality as much as possible.

The participants request an overview of the risks. They want to know who, or what, is initiating them and which assets they harm. The analysis secretary models the risks with their associated risk values in a risk diagram according to the guidelines (see summary). The final risk diagram for unwanted incidents accidentally caused by employees is shown in Fig 13. Since the risk of compromising the confidentiality of health records is within the acceptable risk levels it will not be assessed in the treatment identification.

Step 6 — summary

Tasks:

- likelihood and consequence estimates should be confirmed or adjusted,
- the final adjustments of the acceptable area in the risk matrices should be made (if needed),
- an overview of the risk may be given in a risk diagram.

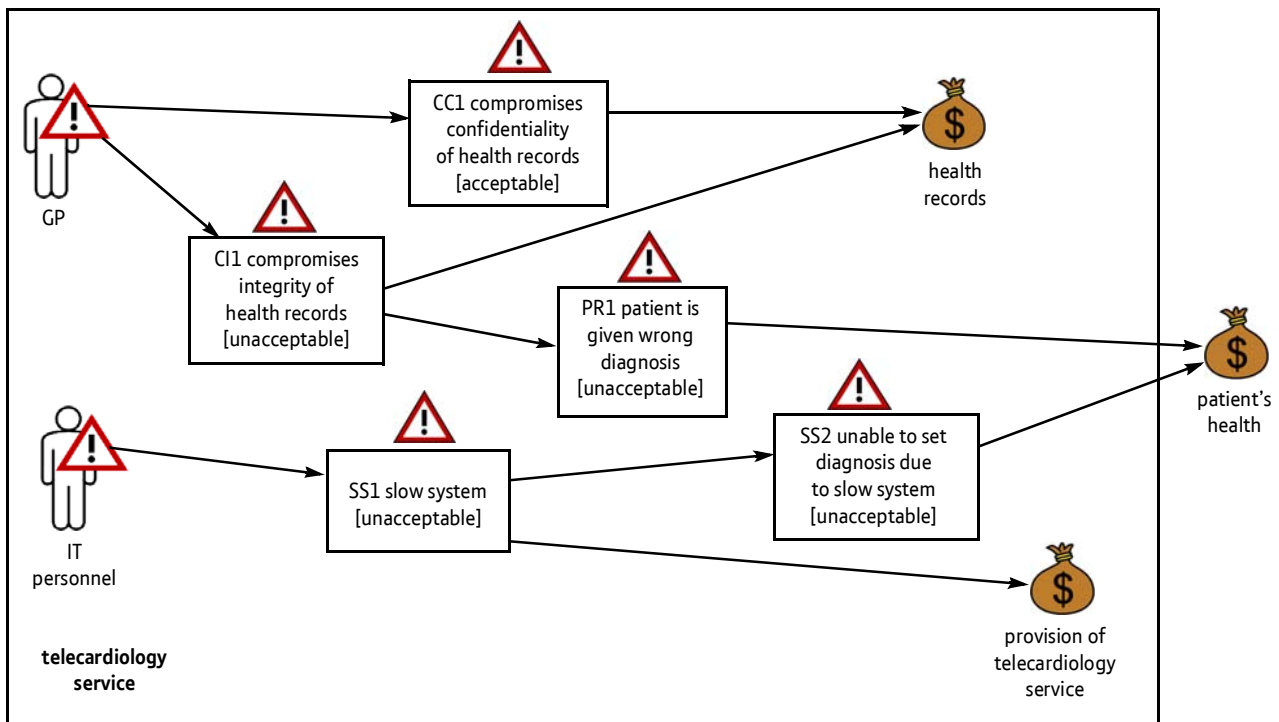


Fig 13 Risk overview.

People that should be present:

- security analysis leader (required),
- security analysis secretary (required),
- representatives of the client:
 - decision makers (required),
 - technical expertise regarding the target (required/optional⁴),
 - users (required/optional⁴).

Modelling guideline:

- risk diagrams:
 - use the threat diagram and replace all unwanted incidents with risk symbols, showing a short risk description and whether the risk is acceptable or not,
 - remove threat scenarios and vulnerabilities, but keep the relations between the threats and the risks,
 - if useful, split the risk diagrams into several diagrams according to type of threat, part of the target or asset importance (e.g. show all risks related to network, all risks for specific assets).

⁴ Depending on whether this step is included in step 7 or not — if it is part of step 7's workshop, all representatives should be present, otherwise it may be sufficient for only the decision makers to be present.

8. Step 7 — risk treatment

The last step in the security analysis is the treatment identification, which is also often organised as a workshop. The risks that are not acceptable are all evaluated in order to find means to reduce their likelihood and/or consequence. Since treatments can be costly, they are assessed with respect to their cost/benefit, before a final treatment plan is made.

The initial treatment diagrams are similar to the final threat diagrams except that every relation between an unwanted incident and an asset representing an unacceptable risk is symbolised with a risk icon and an identifier.

The analysis leader presents each of the threat diagrams showing the unacceptable risks. He knows that analysis participants often find it most intuitive to address vulnerabilities when looking for treatments. Hence, he highlights the possibility of treating other parts of the target as well, such as threats or threat scenarios. The participants become involved in a discussion of potential treatments, and decide which ones will reduce the risks to acceptable levels. On some occasions, if focus is slightly out of scope, the analysis leader suggests treatments taken from best-practice descriptions for network solutions, encryption, etc. to help the discussion back on track. The diagrams are annotated with the identified treatment options indicating where they will be implemented. Finally, the following treatments are suggested and annotated to the treatment diagram in Fig 14:

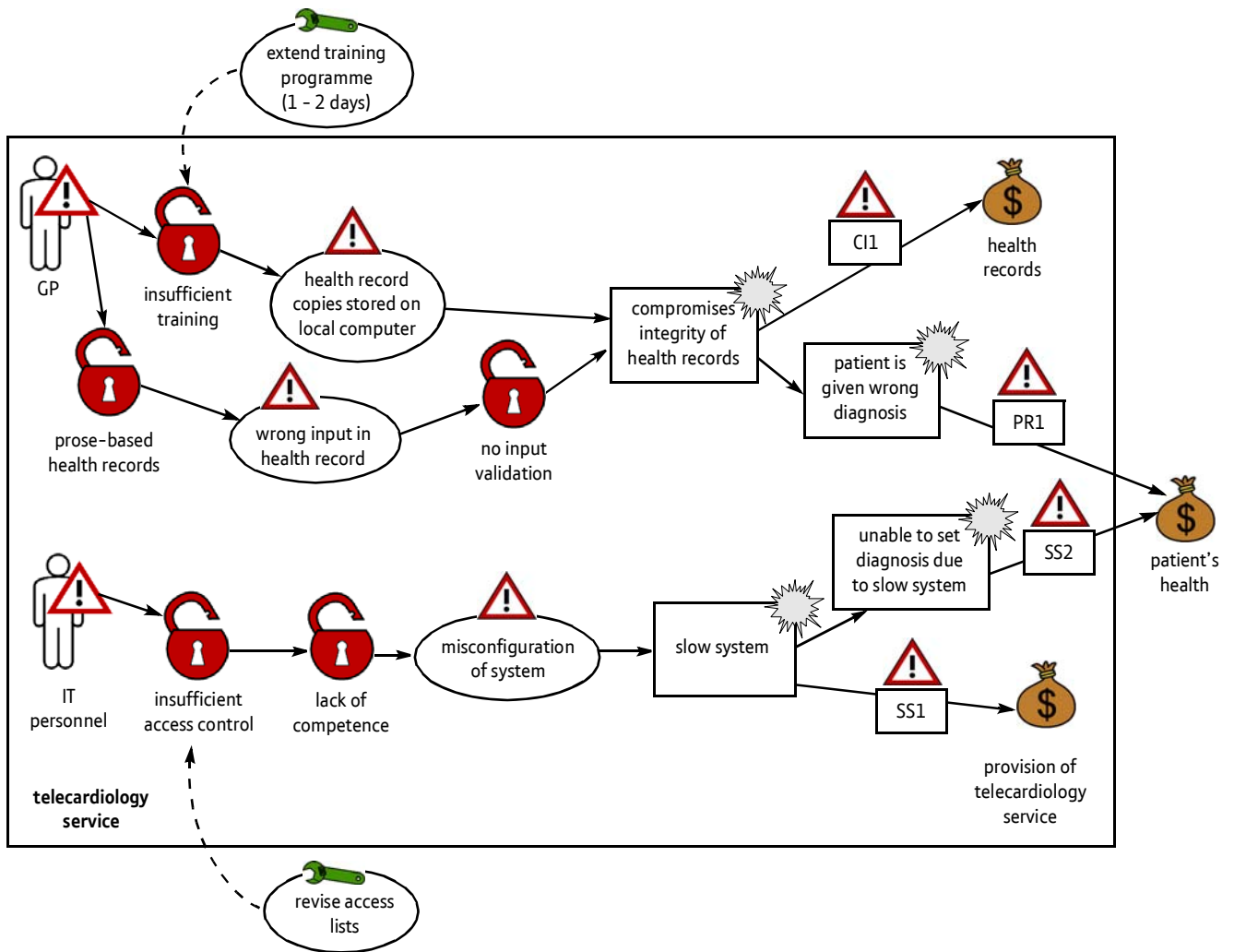


Fig 14 Treatment diagram.

- extend the training programme for practitioners by 1-2 days, with a special focus on security aspects,
- revise the list of people that have maintenance access, and restrict access to only the users that have competence on critical configuration tasks.

When the final results from the analysis are to be presented to the client and other interested parties, an overview of the risks and the proposed treatments is useful. In our case the treatment overview diagram of Fig 15 is used for this purpose.

Step 7 — summary

Tasks:

- add treatments to threat diagrams,
- estimate the cost/benefit of each treatment and decide which ones to use,
- show treatments in risk overview diagrams.

People that should be present:

- security analysis leader (required),
- security analysis secretary (required),
- representatives of the client:
 - decision makers (required),
 - technical expertise (required),
 - users (required).

Modelling guidelines:

- treatment diagrams:
 - use the threat diagrams as a basis and annotate all arrows from unwanted incidents to assets with risk icons, showing only the unacceptable risks,

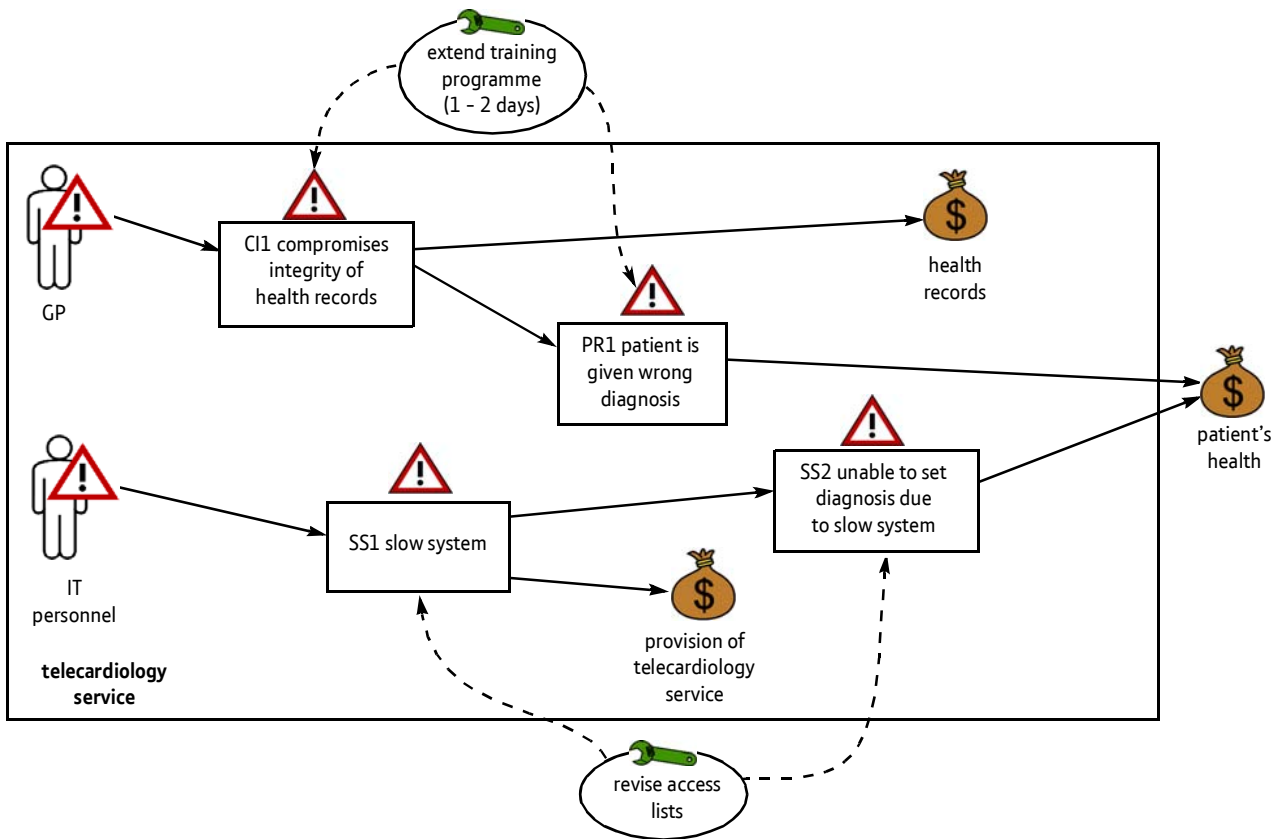


Fig 15 Treatment overview diagram.

— annotate the diagram with treatments, pointing to where they will be applied,

- treatment overview diagrams:

— using the risk diagrams as a basis, remove the acceptable risks,

— add treatments according to the treatment diagram(s).

- risk diagrams to show an overview of which risks are acceptable and which are not — the overview also shows which threats are involved and which assets they may harm,

- treatment diagrams to capture treatment suggestions for the unacceptable risks — the final treatments are also shown in a treatment overview diagram suitable for presentations.

9. Conclusions and related work

This paper has demonstrated the use of the CORAS method in the security analysis of a telemedicine system. The main focus has been on the CORAS security risk modelling language and its role as a medium for communication and interaction during the various steps. In particular, the following core CORAS diagrams have been used:

- asset diagrams to capture and relate the main assets to be protected,
- threat diagrams to show the most relevant threats, e.g. to investigate how employees may accidentally cause risks — the threat diagrams have also been used to capture likelihood and consequence estimates for each risk,

This paper has by no means presented all aspects of CORAS. In particular, the CORAS security risk analysis and modelling tool has not been covered.

The CORAS method and language have been developed iteratively based on experiences from numerous industrial field trials and empirical investigations. The field trials were carried out within the setting of the research project SECURIS (152839/220). Each field trial (in the form of a security analysis) was set to about 250 person hours from the analysts and 50-100 hours from the client. The clients and scopes for the analyses were:

- 2003: NetCom (telecom), information sharing service [7],
- 2004: DNV (vessel classification), information sharing service,

- 2004: NetCom (telecom), mobile access to personal information,
- 2005: Statnett (energy), control and supervisory system,
- 2005: Hydro Aluminium (metal production), control and supervisory system,
- 2006: Hydro (corporate), document management system.

A number of empirical investigations have also been carried out in order to improve the method and the language. These included:

- an experiment on the use of special icons in a modelling language, and how it helps people understand the models better — this experiment also included an initial exploration of the language's conceptual foundation using master students in software engineering [8],
- a survey of security risk analysis terminology that helped us to identify the intuitive and more problematic parts of the language's conceptual foundation — the subjects were professionals and master students within software engineering [9],
- an experiment that identified the preferred modelling alternatives among professionals within software and security [10].

The CORAS method follows the structure of the Australian/New Zealand Standard for Risk Management [4]. The originality of CORAS lies in the strong focus on graphical models during structured brainstorming, both for risk identification and documentation. Compared to two major alternative risk analysis methods CRAMM [11] and OCTAVE [12], the model orientation, instead of just text and table-based documentation, is the main difference.

The structured brainstorming activity in CORAS is inspired by HazOp [13], but developed further to make use of system models as input to the analysis, and threat and risk models as support during the analysis and documentation of outcome.

Misuse cases [14—16], a variation on UML use cases [1, 17] that characterises what a system should not allow, were an important source of inspiration in the development of the CORAS language. The first version of the CORAS language was in the form of a UML profile [18, 19] (i.e. a specialisation of UML) based on use cases. As a result of our work to satisfy the modelling needs in security risk analyses, it has evolved into a more specialised and refined language. Other UML-based and security-oriented languages exist, e.g. UMLsec [20] and SecureUML [21], but they focus on specifying security properties of systems rather than documenting security risks.

Various tree-based notations that support security risk analysis, especially for calculating likelihood, have been around for several years. Most notable are fault trees [6] (and variants like attack trees [22] and threat trees [23]) and event trees [24]. Fault trees are based on decomposing an unwanted incident into its sub-events and use this information to calculate the probability of the top event. With event trees the aim is to model how an unwanted incident may propagate and find the probability and consequence of events originating from the initial event. Cause-consequence diagrams [25] may be seen as the combination of fault trees and event trees. The CORAS diagrams can be said to be inspired by these notations, but do not require input in terms of exact probabilities.

What characterises CORAS is the focus on security risk of information systems, the heavy use of models to guide and structure the analysis, and the specialised language for documenting and communicating intermediate as well as the final results of the analysis. In addition, CORAS comes with a computerised tool that fully integrates all the steps of the analysis from initial planning to production of final reports.

Acknowledgements

The CORAS method has mainly been developed in the 5th Framework IST project CORAS (IST-2000-25031) and the Research Council of Norway projects SECURIS (152839/220) and COBRA (152209/431). Some of the work has also been funded by 6th Framework IST projects TrustCoM (01945) and MODELWARE (511731). The authors acknowledge and thank collaboration partners within these projects for stimulating input and feedback since the start-up of the CORAS project in 2001. The authors have also benefited considerably from teaching this material at the University of Oslo and receiving the students' responses.

References

- 1 OMG: 'Unified Modeling Language (UML): Superstructure, version 2.0', Object Management Group (2005).
- 2 Lund M S, den Braber F, Stølen K and Vraalsen F: 'A UML profile for the identification and analysis of security risks during structured brainstorming', SINTEF ICT, Tech report STF40 A03067 (2004).
- 3 Stathiakis N, Chronaki C, Skipenes E, Henriksen E, Charalambous E, Sykianakis A, Vrouchos G, Antonakis N, Tsiknakis M and Orphanoudakis S: 'Risk assessment of a cardiology eHealth service in HYGEIAnet', in Proc Computers in Cardiology (CIC'2003), pp 201—204 (2003).
- 4 AS/NZS4360, Australian/New Zealand Standard for Risk Management: Standards Australia/Standards New Zealand (2004).
- 5 ISO/IEC13335, Information Technology — Guidelines for management of IT Security (1996—2000).
- 6 IEC61025, Fault Tree Analysis (FTA) (1990).
- 7 den Braber F, Mildal A B, Nes J, Stølen K and Vraalsen F: 'Experiences from using the CORAS methodology to analyse a Web application', Journal of Cases on Information Technology, 7, No 3, pp 110—130 (2005).

- 8 Hogganvik I and Stølen K: 'On the comprehension of security risk scenarios', in Proc Int Workshop on Program Comprehension (IWPC'05), pp 115—124 (2005).
- 9 Hogganvik I and Stølen K: 'Risk analysis terminology for IT-systems, does it match intuition?' in Proc Int Symposium on Empirical Software Engineering (ISESE'05), pp 13—23 (2005).
- 10 Hogganvik I and Stølen K: 'A graphical approach to risk identification, motivated by empirical investigations', in Proc MoDELS'06 LNCS 4199, pp 574—588 (2006).
- 11 Barber B and Davey J: 'The Use of the CCTA Risk Analysis and Management Methodology CRAMM in Health Information Systems', in Proc MEDINFO'92, pp 1589—1593 (1992).
- 12 Alberts C J and Dorofee A J: 'OCTAVE Criteria Version 2.0', Tech report CMU/SEI-2001-TR-016, ESC-TR-2001-016 (2001).
- 13 Redmill F, Chudleigh M and Catmur J: 'HAZOP and Software HAZOP', Wiley (1999).
- 14 Alexander I: 'Misuse cases: Use cases with hostile intent', IEEE Software, 20, No 1, pp 58—66 (2003).
- 15 Sindre G and Opdahl A L: 'Eliciting Security Requirements by Misuse Cases', in Proc TOOLS-PACIFIC, pp 120—131 (2000).
- 16 Sindre G and Opdahl A L: 'Templates for Misuse Case Description', in Proc Workshop of Requirements Engineering: Foundation of Software Quality (REFSQ'01), pp 125—136 (2001).
- 17 Jacobson I, Christerson M, Jonsson P and Övergaard G: 'Object-Oriented Software Engineering: A Use Case Driven Approach', Addison-Wesley (1992).
- 18 Lund M S, Hogganvik I, Seehusen F and Stølen K: 'UML profile for security assessment', SINTEF Telecom and Informats, Tech report STF40 A03066 (2003).
- 19 OMG: 'UML Profile for Modeling Quality of Service and Fault Tolerance Characteristics and Mechanisms', Object Management Group 2006.
- 20 Jürjens J: 'Secure systems development with UML', Springer (2005).
- 21 Lodderstedt T, Basin D and Doser J: 'SecureUML: A UML-Based Modeling Language for Model-Driven Security', in Proc UML'02, LNCS 2460, pp 426—441 (2002).
- 22 Schneider B: 'Attack trees: Modeling security threats', Dr Dobb's Journal, 24, No 12, pp 21—29 (1999).
- 23 Howard M and LeBlanc D: 'Writing Secure Code', 2nd edition, Microsoft Press (2003).
- 24 IEC60300-3-9, Event Tree Analysis in Dependability management — Part 3: Application guide — Section 9: Risk analysis of technological systems (1995).
- 25 Nielsen D S: 'The Cause/Consequence Diagram Method as a Basis for Quantitative Accident Analysis', Danish Atomic Energy Commission, RISO-M-1374 (1971).



Folker den Braber holds a Master's degree in Computer Science from the University of Leiden, The Netherlands.

He has been working as a research scientist at SINTEF Information and Communication Technology since 2001.

His main research interests are in semi-formal modelling, system architecture and development, and security analysis and architecture.



Ida Hogganvik received her Master's degree in Computer Science from the Norwegian University of Science and Technology in 2003. She is currently working on her PhD at the University of Oslo/SINTEF Information and Communication Technology.

Her field of interests includes graphical modelling in security analysis, empirical research methods and information visualisation techniques.



Mass Soldal Lund holds a Master's degree in Computer Science from the University of Oslo, specialising in formal methods.

Currently he is employed as a research scientist at SINTEF Information and Communication Technology, where he has been working with risk analysis and threat modelling since 2001, and at the University of Oslo, where he is studying for his PhD.



Ketil Stølen is Chief Scientist and Group Leader at SINTEF Information and Communication Technology. He received his PhD from Manchester University and since 1998 he has been a Professor in computer science at the University of Oslo. Earlier employers include Munich University of Technology and the OECD Halden Reactor Project.

He was the technical manager of the EU-project CORAS running from 2001 to 2003.



Fredrik Vraalsen received his Master's degree in Computer Science from the University of Illinois at Urbana-Champaign in 2001. Until 2006 he was employed as research scientist at SINTEF Information and Communication Technology, working on mobile systems and technologies, systems architecture and design, security risk analysis, and system development.

He currently works as senior developer of J2EE solutions at Unified Consulting.