An Investigation into Error Detection and Recovery in UK National Health Service Screening Programmes

Nick Chozos

Thesis submitted for the
Degree of Doctor of Philosophy (Ph.D)

Department of Computing Science

August 2009

Abstract

The purpose of this thesis is to gain an understanding of the problems that may impede detection and recovery of NHS laboratory screening errors. This is done by developing an accident analysis technique that isolates and further analyzes error handling activities, and applying it in four case studies; four recent incidents where laboratory errors in NHS screening programmes resulted in multiple misdiagnoses over months or even years. These errors resulted in false yet plausible test results, thus being masked and almost impossible to detect in isolated cases.

This technique is based on a theoretical framework that draws upon cognitive science and systems engineering, in order to explore the impact of the plausibility of false test results on the entire process of error recovery. The four analyses are then integrated and compared, in order to produce a set of conclusions and recommendations.

The main output of this work is the "Screening Error Recovery Model"; a model which captures and illustrates the different kinds of activities that took place during the organizational incident responses of these four incidents. The model can be used to analyze and design error recovery procedures in complex, inter-organizational settings, such as the NHS, and its Primary/Secondary care structure.

Thesis statement

This thesis aims to contribute to the safety and overall quality of screening programmes in the NHS, by enhancing our understanding of the problems that may impede detection and recovery of screening errors.

Table of Contents

List of figures

List of tables

Acknowledgements

There are many people I would like to thank for making this study not only possible, but also very enjoyable. I would like to begin with my supervisor, Chris W. Johnson who has been very motivating, supportive and understanding through these years.

I would also like to thank Dr. James Barnes and all staff at Crookston medical practice for welcoming me in their workplace, taking interest in my work and openly sharing their experiences which were very insightful. In a similar manner, many thanks go to Dr. Frank Finley at the Biochemistry Department at the Northern General hospital in Glasgow.

Furthermore, I am very grateful to Dr. Robert L. Wears and staff in the Department of Emergency Medicine at the University of Florida, who I was lucky to visit during November and December 2006.

I would also like to thank Philip Gray and Elizabeth McFarlane at Lilibank Gardens, and Philipe Palanque and Sandra Basnyat at Toulouse.

Finally, many thanks go to the new friends I made in Glasgow, which is now my second home.

This work is dedicated to my parents and my sister who I very much missed during the three years I spent in Glasgow.

This study was carried out in the Department of Computing Science of the University of Glasgow during the years of 2004-2008, and was funded by the European Commission's ADVISES Research Training Network. Part of this work was also carried out at Université Paul Sabatier in Toulouse, and the University of Florida, thanks to the 'Ken Browning' travelling scholarship in Computing and Medicine.

Declaration of ownership

I hereby declare that I composed this thesis myself, and that it describes my own work. It has not been submitted for examination at this or any other academic institution until this time.

Nick Chozos
London, England
August 2009.

Chapter 1: Introduction

Laboratory and radiology departments are playing an increasingly critical role in modern healthcare. In the past 25 years, advances in medical knowledge and technology have created the opportunity for better and faster patient diagnosis; for instance, with new high-speed analyzers laboratory testing can be largely automated, while with an information-technology based infrastructure, specialist doctors can now perform the interpretation of X-rays from a distance, without the patient having to go to the hospital [Brennan, 2005].

The benefits of these innovations can be seen in the Breast Cancer Screening Programme of the UK's National Health Service (NHS). Since its introduction in 1987, the programme resulted to a 25% drop in mortality rates attributed to breast cancer by 2000 [NHS Advisory Committee for Breast Cancer Screening, 2006]. Similar success has been achieved by other NHS screening programs, altogether contributing significantly to the timely diagnosis of various forms of cancer, Sexually Transmitted Infections (STI), and other critical conditions.

Despite the significant advances in laboratory medicine, hospitals and laboratories still remain concerned about the accuracy, validity and reliability of clinical test results

[Plebani and Carraro, 1997; Hickner et al., 2006; Schiff, 2006]. As several challenges and problems remain in the already complex setting of diagnostic networks (e.g., the definition of an acceptable error range, delays in following-up critical test results), the drastic changes that information technology has brought about have created the potential for new kinds of error, which—although rare—can be significantly more disastrous in extent.

Table 1.1 summarizes three serious screening incidents which had multiple adverse outcomes over long periods of time. The table below is based on the subsequent inquiry reports that were produced [Ferres et al., 2001; Commission for health improvement, 2002; Baker, 2006]; immediate cause refers to the cause that initiated the incident, while incident-prolonging causes are issues that resulted in not detecting the immediate cause and/or poorly handing the incident. Similar screening incidents have occurred in the UK [The Guardian, 2006], USA [Wears, 2003] and Canada [Bernstein, 2003].

Incident	Investigation findings	Patients affected
1. ***Down's syndrome screening errors***, Sheffield Northern General Hospital, Immunology Dept. *January-May 2001*	***Immediate cause:*** A software bug was affecting an algorithm used to calculate the likelihood of pregnant women giving birth to children with Down's ***Incident-prolonging factors*:** User interface deficiencies, audits not carried out as planned, incident log books not used as specified, poor communication between staff groups	- *158 high-risk pregnancies missed* - 2 women had late abortion - 2 women gave birth to children with Down's
2. ***Breast cancer screening errors***, Hammersmith London, Breast Cancer screening service *1993-October 2000*	***Immediate cause:*** Confusing notation for denoting positive and negative results ***Incident-prolonging factors:*** No robust protocol for ensuring women received the correct result, strained relationships between staff groups, poor handling of complaints	- *Over 12,000 incorrect tests* - 17 patients more critical - 1 death
3. ***Breast cancer screening errors,*** Manchester, Breast Cancer screening service *April 2003–January 2006*	***Immediate cause:*** 'Human' error of a single radiologist, who misinterpreted multiple mammograms ***Incident-prolonging factors:*** No double-checking of radiology reports which is common practice, lack of safeguards	- *176 mammograms misinterpreted* - 28 cancers missed, out of which 17 were very critical

Table 1.1: NHS screening incidents

Although the immediate causes that led to these unfortunate events vary (i.e., hardware and software bugs, problematic notation for denoting positives, human error), all of these incidents were prolonged for several months by relatively common organizational problems (e.g., communication breakdowns, lack of safeguards, poor handling of complaints); however, the most important aspect of these failures is that when the errors manifested, false test results were *plausibly acceptable*, masking errors and allowing for them to be used in the diagnostic process. Detection was consequently only possible over time, when experienced staff became increasingly alarmed over a lack of positive results reported from the laboratory.

The study presented in this thesis is therefore an attempt to identify and analyze the factors that inhibit laboratory error detection and recovery. While there has been considerable work in the study of laboratory error, little attention has been paid to the impact of such errors to healthcare systems overall and how errors are detected and dealt with [Plebani and Carraro, 2004]. In addition, this work may be distinguished from previous error handling studies by incorporating the concept of *problem detection* in the overall *error recovery process*. Problem detection refers to the concerns over a potential error, as opposed to error detection, which is the identification of an error) [Klein et al., 2005].

In order to achieve these purposes, an *accident analysis* tool has been developed which focuses on the error handling activities that took place during an incident. This tool is the primary contribution of this thesis, which has been specifically tailored to help analyze these healthcare events. The development of such a tool was found necessary as existing accident analysis approaches do not take a structured perspective on the sequence of events that form an error recovery process. The accident analysis approach suggested has been used to analyze four incidents; resulting findings were then integrated and compared in order to draw high level conclusions about the factors that limit the ability of healthcare systems to detect, control and recover from laboratory error.

The purpose of this chapter is to discuss the aims and objectives of the study, and present an overview of the thesis.

1.1 Aims and objectives

The work presented in this thesis focuses on the events that take place once there are initial concerns that 'something is wrong', but does not examine the causal factors, i.e., what led to the errors in the first place. The high-level goal of this study is to understand how errors could be better detected, contained and controlled, in order to help healthcare professionals limit the consequences on human life to the smallest extent.

The aims of this thesis are the following:

- ***Primary Aim:*** **To gain a detailed understanding of the factors that affect detection and recovery of screening errors**

  Investigations into the incidents discussed in Table 1.1 produced detailed conclusions and recommendations regarding the improvement of organizational response to errors. The primary aim of this thesis is to utilize these findings by comparing and integrating them, and further analyzing them with a scientific method that can be useful for policy-makers, system and medical device designers. The results of this analysis will be presented in Chapter 6, which will conclude with a detailed model that presents all the activities that may take place during laboratory error handling.

- ***Secondary Aim:*** **To generate recommendations for the improvement of laboratory error handling**

  The basis for these recommendations will be the model that has been developed in order to meet the Primary Aim. Recommendations will focus on improving each stage of laboratory error recovery. An important aspect of these recommendations

is that they are based on the relationship between the different stages of error recovery (these will be discussed in detail in Chapter 3, but they are error detection, error indication, further investigation, error explanation, and error correction), e.g. different recommendations regarding incident reporting that has been stimulated from different kinds of detection.

In order to meet these aims, it was necessary to achieve the following research objective:

- ***Research Objective:*** **To develop and validate an accident analysis tool that can be used to identify and analyze error handling activities**

  This accident analysis technique is an adaptation of 'Sequentially Timed Events Plotting' (STEP) [Henrick and Benner, 1983] that has been integrated with error recovery theory in order to take a focus on error handling. The development of Error Recovery-STEP (ER-STEP) will be discussed in detail in *Chapter 4: Research Methodology.*

The following section presents an overview of the methodology used to meet these objectives, and provides an introduction to the findings of this thesis.

1.2 Overview of research methodology and results

Figure 1.1 summarizes the three parts of the methodology and their relationship to the aims and the research objective of this thesis. Step 1 is the development of ER-STEP, Step 3 is the application of ER-STEP for the analysis of four case studies (which were discussed in the introduction of this Chapter) and Step 3 is the integration of the individual findings in order to draw high level conclusions.

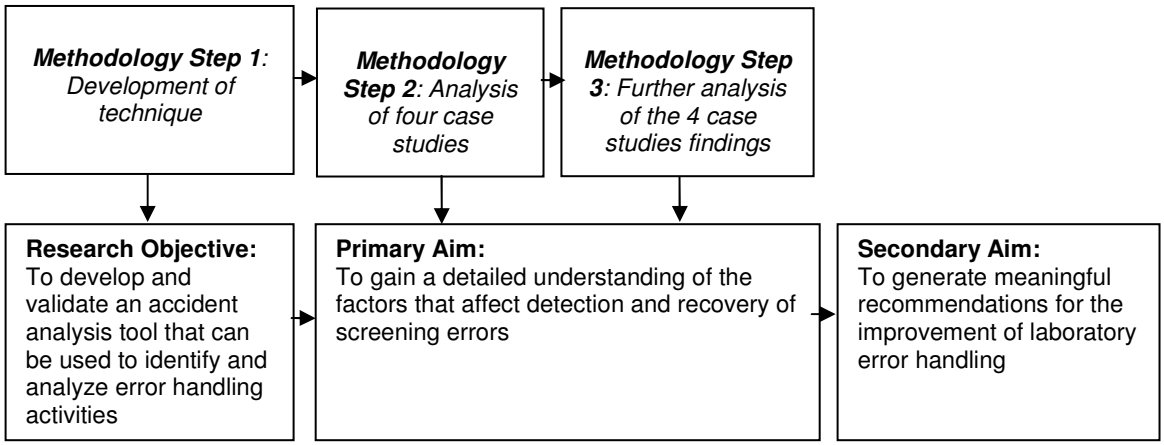


Figure 1.1: Overview of research methodology, aims and objectives.

The analysis of the four incidents resulted in an informed model that describes the various kinds of activities within a healthcare system that make up the organizational response towards the control and correction of a laboratory error.

The model is presented in Figure 1.2 below; it illustrates the relationship between error recovery activities, and can be used to design error recovery processes based on different kinds of detection, including incident reporting schemes and user interface design.

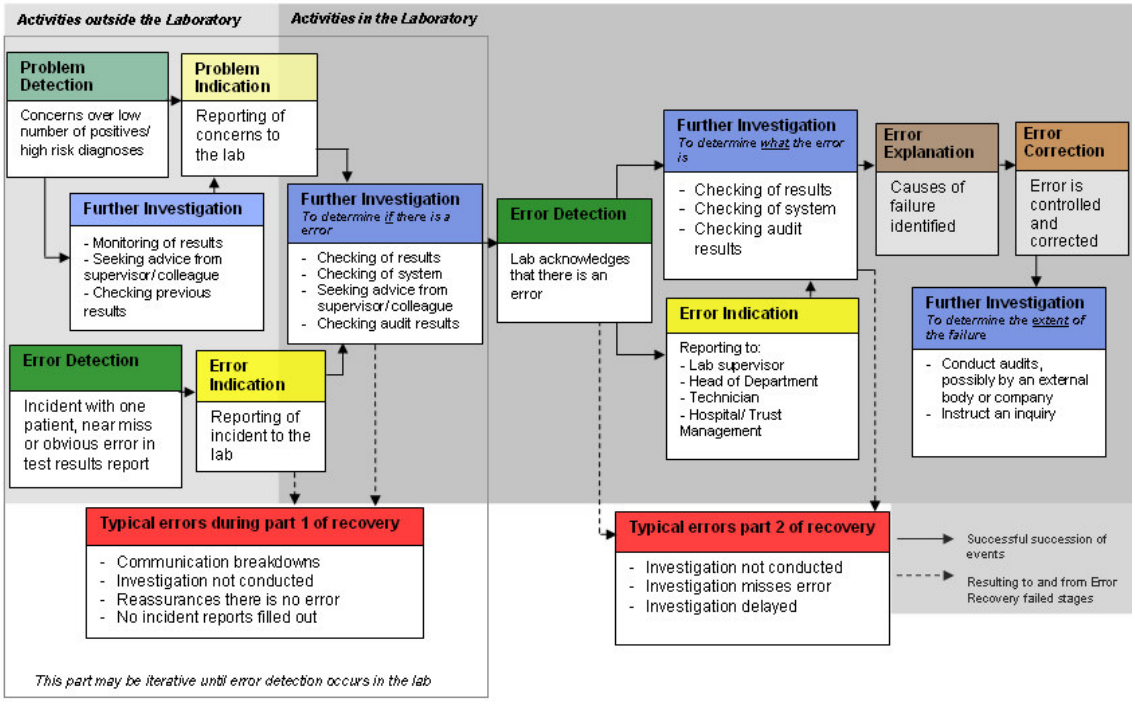


Figure 1.2: The Screening error recovery model.

The model and its development will be explained fully in *Chapter 6: Overview of findings*. This is later used to guide the recommendations that are discussed in *Chapter 7: Recommendations.*

The following section summarizes the contents of each chapter of this thesis.

1.3 Thesis breakdown

Chapter 2 presents an overview of laboratory work and laboratory error, which are the field and the focus of this thesis. First, the organizational structure of NHS diagnostic networks will be presented. This is important as there are several inter-organizational issues that need to be considered. This chapter will also discuss several studies that have attempted to identify and classify laboratory and radiology error types and frequencies. This chapter will conclude with a definition of the problem that this thesis aims to tackle.

Chapter 3 contains the literature review that this thesis draws upon. Theory on problem detection and error detection, error recovery strategies and models, and error management will be presented and discussed. This chapter will conclude with a new Error Recovery Framework, which has been developed for the purposes of this thesis. This framework builds upon the literature that has been reviewed but extends it by incorporating problem detection in the error recovery process.

Chapter 4 describes the research methodology that has been used. This chapter discusses some existing approaches to accident analysis and elaborates on the need for an error-handling focused analytical approach. This chapter will conclude with a presentation of the accident analysis tool that has been developed for the purposes of this thesis (Research Objective).

Chapter 5 presents the analysis of four case studies in detail. These are the three incidents summarized in Table 1.1, as well as one incident that took place in the USA. Each case study, along with graphical illustrations of the applied analytical technique will be presented in this chapter.

Chapter 6 presents an overview of the findings of the four analyses. This chapter will conclude with a model that describes the 'Laboratory Error Handling Process', which is an overview of the error recovery activities that may take place within a diagnostic network. This is where the Primary Aim of this thesis has been met.

Chapter 7 discusses some preliminary recommendations for the improvement of laboratory error detection and recovery (Secondary Aim). These recommendations will consider improvements, interventions and new ways for better dealing with laboratory error.

Chapter 8 presents the validation of ER-STEP. The method undertaken for validation and the subsequent results are discussed here.

Finally, *Chapter 9* contains the overall conclusions that have arisen from this work, and discusses some possible directions for further research. This chapter is followed by a list of references and two appendices.

Chapter 2: Field and focus

The previous chapter introduced the aims and objectives that this thesis hopes to achieve. This chapter will present the *field and focus* of this study. An overview of laboratory medicine and screening programmes in the NHS will first be presented. This discussion will proceed with a description of the inter-organizational networks that take part in diagnostic services, focusing on the dependencies that are developed upon laboratory and radiology departments within NHS Trusts, as an error in one laboratory can propagate in various organizations. This work, therefore, takes place in the *field* of diagnostic networks in the NHS, laboratory medicine, and the technological and procedural aspects that support such networks.

The *focus* of this thesis is laboratory error, and in particular in screening services. There is much ongoing work aiming at the analysis of laboratory error [e.g. De Boer et al., 2002; Sirota, 2005; Frable, 2006]; however, these studies tend to be confined within a specific laboratory, taking a rather quantitative approach towards the measurement of error types' frequencies. It has also been suggested that most laboratory error studies do not consider the impact such errors have on patients, as laboratories do not maintain information about the results of their work in terms of patient outcomes [Bonini et al., 2002].

This chapter will provide some background in laboratory medicine services and in the UK NHS in particular, focusing on screening services, before discussing laboratory error and recent work that identifies types and frequencies of error that can take place in a laboratory setting. This information will then be related to the incidents presented in Table 1.1 in order to place the problems that this thesis aims to tackle within the wider context of healthcare systems.

2.1 Laboratory medicine

Laboratory and radiology services are an integral part of diagnosis and monitoring of patients. On a daily basis, pathology, biochemistry, microbiology, immunology and other types of laboratories produce a wide range of test reports which are used to support clinicians' decision upon patient treatment [Brennan, 2005]. Overall, there are three types of laboratory testing:

1. **Screening test**: a test in search of a disease in a person who does not appear to have it; e.g. PSA test (Prostate Specific Antigen) for prostate cancer.
2. **Diagnostic test**: a test for a specific, particular disease; e.g. lung cancer.
3. **Monitoring test:** a test which helps doctors keep track of how a patient is doing with a known disease; e.g. monitoring a diabetic patient.

The testing process is made up of three stages: *Pre-analytic, analytic* and *post-analytic*. The pre-analytic phase is structured around the ordering and implementation of the test. The analytic phase conducts the specimen analysis. Post-analytic is the communication,

documentation and usage of the test results [Sirota, 2005]. Figure 2.1 illustrates an 'error free' testing process.

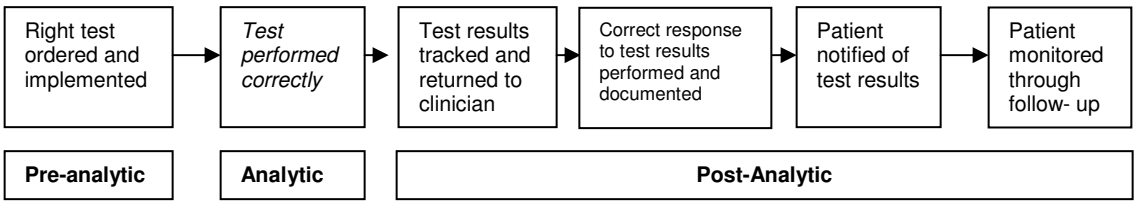


Figure 2.1: An 'error-free' testing process [adapted from Hickner et al., 2005].

The following section will discuss how diagnostic services are provided by the NHS.

2.2 NHS diagnostic services

The NHS can be seen to be divided in two sections: Primary and Secondary care. Primary care is the 'frontline' service, which is the first point of contact for patients. Primary care consists mainly of General Practice (GP) clinics and surgeries, as well as dentists, opticians and pharmacists. When a patient walks in the GP practice, a prognosis or initial consultation may conclude that laboratory testing is required, which will be carried out in a hospital (Secondary care). Specimens are taken at the GP premises and are then sent to the laboratory[1]. As several hospitals, departments and GPs are attached to one laboratory, several specimens are analyzed in a routine, batch process, and the results are then sent back to each GP.

Figure 2.2 illustrates a typical network of primary and secondary care organizations which are all dependent on a single laboratory for the provision of diagnostic services.

[1] More information can be found at www.nhs.uk/conditions/nhs/Pages/Definition.aspx?url=Pages/what-is-it.aspx, last accessed 05-Oct-08

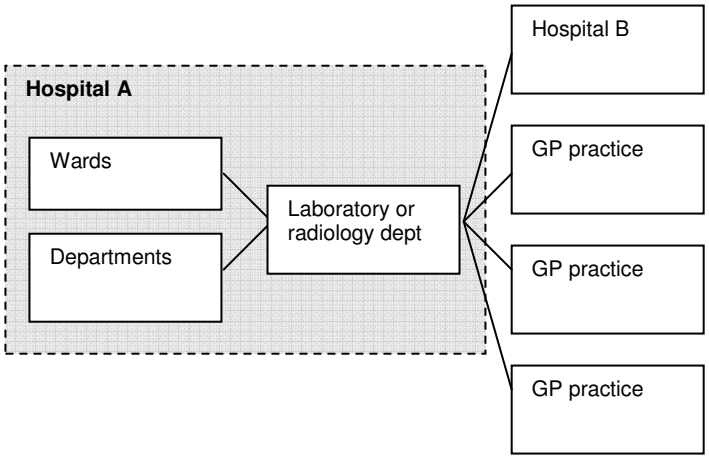


Figure 2.2: NHS diagnostic networks structure.

The UK model of diagnostic services provision is somewhat different from the one in the USA, where there is an attempt to "*bring lab-testing closer to the patient*". In the USA, testing facilities may also be offered at the first point of contact (for the UK in GPs or walk-in centres). When the patient goes to a clinic, there will first be a battery of common tests before the doctor even sees the patient, which will then be taken into account along with patient history, whereas in the UK such tests have to be requested following a consultation and sent back—a process which will may take days to perform [Brennan, 2005]. In this way, the NHS model of laboratory services is significantly different from the USA model.

Nonetheless—and as we shall see later on—the NHS diagnostic model has seen notable improvements, with important cost reduction, increases in productivity and offers more diagnostic services. Much of the success of this model is based on the utilisation of modern technologies. The next section will briefly discuss some of the advances that

have driven much of the drastic structural and workflow changes seen in the NHS in the past twenty years.

2.2.1 Laboratory information systems

While the number of tests ordered has increased substantially over the past decade, laboratory systems also steadily growing to offer more critical diagnostic services [Smith and McNeely, 1999; Schiff, 2006]. In order to cope with these increasing demands, the testing process is utilizing a combination of complex technologies that can automate the analytic stage. This combination involves hardware and software that is used for specimen analysis, and the subsequent calculations that need to be performed in order to derive the requested test results.

In addition, information technology applications have also been introduced for the request, archiving and communication of test results and radiology reports. For instance, Picture Archiving and Communication Systems (PACS) are digital imaging solutions that can also distribute X-ray films over a computer network. In the near future, all NHS Trusts will have a PACS system [Brennan, 2005]. NHS Scotland has introduced the Electronic Clinical Communications Implementation (ECCI), which aims at facilitating communications between primary and secondary care for the request and follow-up of laboratory tests, patient referrals, outpatient appointment etc. [Pagliari et al., 2004].

These innovations have significantly reduced costs and created an infrastructure which allows for a more efficient and productive testing process. Laboratories can thus offer more services to more patients, by automating a great part of analytic testing and

communication with GP practices, departments and wards as well as other hospitals. Such advances have formed the basis for the success of screening programmes in the NHS, which are discussed in the next section.

2.2.2 NHS Screening programmes

Screening programmes aim at diagnosing critical conditions such as cancer by routinely evaluating patients that are likely to have that specific condition. For instance, under the NHS Breast Cancer Screening Programme, women aged 50-64 are invited for mammography screening once every three years [Advisory Committee on Breast Screening, 2006].

Breast Cancer Screening was introduced in 1986. Since then, the NHS has grown to offer a variety of screening programmes[2]. They can be summarized as follows:

- **Cancer screening:** Breast cancer, cervical cancer and bowel cancer. There is currently no national screening programme for prostate cancer, but a risk management programme is available.
- **Vascular diseases:** Heart disease, diabetes and stroke.
- **Sexually transmitted infections:** Human Immunodeficiency Virus (HIV), Chlamydia, Hepatitis B and C and others.
- **Screening for pregnant women and/or their newborn babies:** Down's syndrome, fetal anomalies, hearing, hepatitis B and HIV.

[2] More information can be found at www.screening.nhs.uk, last accessed 05-Oct-08

NHS screening programmes have achieved notable success; for instance, Breast Cancer screening has contributed to a 25% reduction in mortality attributed to the malignancy. In the same time, the number of women screened annually has been steadily increasing. By 2005, there were approximately 1.3 million women screened, whereby 10,000 cancers are identified per year. Timely detection has then resulted to, not only a reduction in mortality but also a reduction to the number of mastectomies [Advisory Committee on Breast Screening, 2006].

As mentioned previously, many screening programmes are driven by the capabilities of new technologies, but involve careful consideration of policy and the development of appropriate management structures. New screening programmes are evaluated by the National Screening Committee (NSC) which uses research evidence and the skills of multi-disciplinary expert groups to develop policies for screening. The aim of the NSC is to ensure that *"screening does more good than harm at a reasonable cost"*[3] by assessing new programmes against a set of recognized international standards.

There has been much criticism about the way with which new screening programmes are introduced. For instance, in 1996, an internal report compiled by the NSC characterized NHS screening programmes a "mess". Table 2.1 summarizes some of the problems identified by the NSC at the time [in Programme Director's report, 2005].

[3] National Screening Committee official website, http://www.nsc.nhs.uk/uk_nsc/uk_nsc_main.htm#remit, last accessed 07-Oct-08

- Unknowing variations in policy, including no policy.
- Unknowing variations in practice.
- Absence of standards.
- Absence of performance measurement.
- Patchy training.
- Poor information for women.
- Lack of clear lines of accountability.

Table 2.1: Problems identified in NHS Breast Cancer Screening Programme [Programme Director's report, 2005].

Continuing in this report, the NSC suggested that is essential to develop clear systems of management which are able to deliver the four functions of a quality assurance programme, namely:

1. Minimizing the risks of error.
2. Dealing with errors and adverse events quickly and compassionately.
3. Continual improvement in performance, either by investment of resources, new technology, or process redesign, and
4. Regular re-setting of quality standards.

This thesis relates primarily to function 2 but also to function 3. The statement of function 2 is an acknowledgement of the problems that the NHS has been facing with regards to screening error. The temporal aspect of error recovery is important because a misdiagnosed patient's condition is most likely to deteriorate with time. This thesis can be seen to contribute to function 3 by promoting the development of more efficient laboratory error recovery strategies.

The following section will discuss quality assurance practices in the NHS.

2.2.3 Diagnostic services quality management

Laboratories and radiology departments have performance standards and systems in place for quality control and quality assurance. Reliability cannot be achieved in a clinical laboratory just through the promotion of accuracy in the analytical phase of the testing process; hence, monitoring all steps in laboratory testing in order to detect and correct defects is very important [Witte et al., 1997]. Quality assurance is therefore applied throughout the testing process (see Figure 2.3).

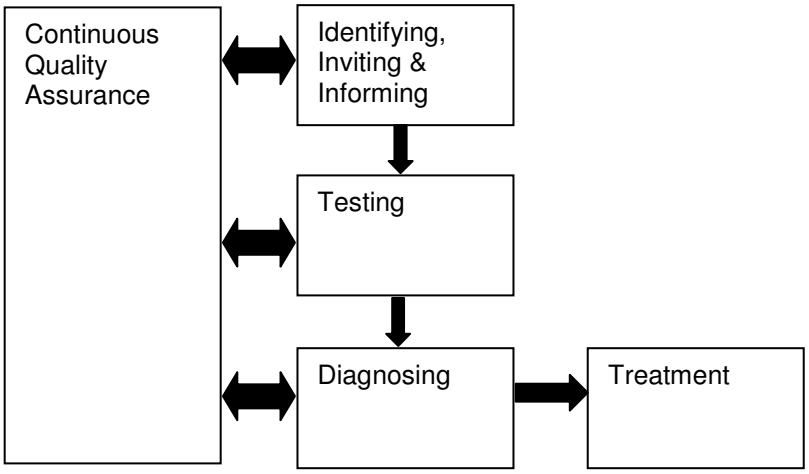


Figure 2.3: Quality assurance in screening services [adapted from Programme Director's report, 2005].

Some of the main aspects of laboratory quality assurance are summarized here:

- **Laboratory standards:** Each diagnostic specialization is governed by guidelines and standards that dictate how laboratories should conduct their services in order to achieve a desirable level of quality. Policy documents cover all aspects of laboratory work, from how testing should be carried out, to the acquisition of medical devices, and algorithms employed in the specific analysis [Johnson and Patnick, 2000].

- **Performance indicators:** Performance indicators are performance standards that are proposed and controlled on a national level by accrediting agencies. These are used to evaluate the performance on all healthcare organizations. Ratings can then be used for benchmarking purposes. Table 2.2 presents 8 performance measures that may be applicable to all clinical laboratories [Howanitz, 2005].

Measure	Laboratory Discipline	Phase of Testing	Frequency of Data Collection
Customer satisfaction	All	All 3	Yearly
Turnaround time	Chemistry, Haematology	All 3	Monthly
Patient Identification	All	Pre-analytic	Monthly
Specimen acceptability	Chemistry, Haematology	Pre-analytic	Monthly
Proficiency testing	All	Analytic	6- 20 specimens per analyte yearly
Critical Value reporting	All	Post-analytic	Monthly
Blood product wastage	Transfusion medicine	Post-analytic	Monthly
Blood Culture contamination	Microbiology	Pre-analytic	Monthly

Table 2.2: Critical laboratory performance measures [taken from Howanitz, 2005].

- **Auditing:** Laboratory audit is concerned with the everyday aspects of the work of the department. Audits are usually organized internally (Internal Quality Control). However, the National External Quality Assessment Service (known as NEQAS) and Clinical Pathology Accreditation schemes can complement the in-house program of audit[4]. The auditors check compliance, non-compliance or possible non-compliance against a checklist and write a report, while the quality system

[4] More information can be found at www.ukneqas.org.uk, last accessed 07-Oct-08

itself can be audited (Figure 2.4 presents a fragment of an example audit checklist[5]). Any faults identified by an audit should lead to immediate corrective action and appropriate changes in documentation, which should be discussed in management reviews.

Laboratory Systems Audit Worksheet

Laboratory:		Date:	
Auditor(s):		Signature:	

Sample Information:

Project Code:		Sampling Date:	
Field ID #:		Lab ID #	

General Information:

	Y/N	Comments
Sample Containers and Equipment Decontamination and Prep:		
Is decontamination procedure acceptable?		
Is sample equipment storage procedure acceptable?		
Is there an existing QC check on bottles and sampling equipment?		
Are certificates for pre-cleaned bottles maintained on file?		
Is the preservation preparation and dispensing documented and traceable?		
Sample Log-in and Receipt:		
When were samples submitted to lab?		
When were samples logged into LIMS?		
Was sample temperature checked upon receipt?		
Was sample preservation checked upon receipt?		
Do LIMS #s match the corresponding field #s?		
Is the sample storage area secured?		
Are samples stored in appropriate refrigerator?		
Was storage refrigerator temperature monitored daily?		

Figure 2.4: Example laboratory audit checklist.

The primary purpose of quality assurance is to ensure that "*the right result on the right specimen of the right patient is accurate, timely and properly interpreted*" [Standards Unit, Evaluation and Standards Laboratory, 2008]. Quality assurance in laboratory systems primarily aims at the avoidance of errors throughout the three phases of the test lifecycle. The next section will introduce laboratory error and discuss various laboratory error related studies that have taken place recently.

[5] Full example laboratory audit checklist can be found at: http://www.evergladesplan.org/pm/recover/recover_docs/wqt/qasr_app_g.pdf, last accessed 07-Oct-08

2.3 Laboratory error

Defining laboratory error is very challenging, as there is much debate regarding acceptable error ranges and to what rate unacceptable test results are tolerable [Blumenthal, 1997]. Not all patients' samples with unacceptable results are equally likely to alter patients' outcomes [Witte et al., 1997]. This confusion makes it very difficult to define laboratory error, and therefore regulate and mandate laboratory quality control practices in laboratory services [Bonini et al., 2002]. Nevertheless, laboratory error can be broadly defined as "*any defect during the entire testing process, from ordering to reporting results*" [Plebani and Carraro, 1997].

From this definition it is obvious that the risks associated with laboratory services are not limited within the premises of the lab and errors in the analytical stage of the testing process. In fact, many misdiagnoses have resulted from requesting the wrong result or mixing up patients, errors occurring during the ordering or use of test results by other clinical units or even institutions [Schiff, 2006].

2.3.1 Error types and frequencies

Table 2.3 summarizes the findings of three studies that have attempted to classify various errors according to the three stages of laboratory testing. These studies were carried out by monitoring different laboratories over periods of several months. Data was derived from a review of test records and audits.

Pre-analytical Phase	Analytical Phase	Post-analytical Phase	Study
- ordering of incorrect test - specimen handling errors - provision of false information to the laboratory	- specimen mix-up - mislabelling of specimens - lack of appropriate measurements - knowledge problems - problematic classification models	- delivery of report to the wrong location - clinician misinterpretation	Sirota, 2005
- failure to order - ordering delay - ordered but not completed - contraindicated	- specimen not sent or not picked up - incorrect preparation - inadequate equipment - results incorrectly processed	- ambiguous report - misread or missed critical report - report not acted on - report information not available when needed	Hickner et al., 2006
- wrong patient name - erroneous specification of hospital unit - physician order missed - order misinterpreted - inappropriate container used	- isolated malfunctioning of instrument - lack of specificity of the method - unacceptable performance	- correction of erroneous finding overlooked - keyboard entry error - turnaround time exceeded - physician not notified of problem	Plebani and Carraro, 1997

Table 2.3: Errors in the testing process.

As we can see, there is a variety of possible errors that can occur during the request, analysis and follow-up of test results. The following table presents the findings of five studies that attempted to measure the frequencies of errors in the three stages of laboratory testing.

Pre-analytical Phase	Analytical Phase	Post-analytical Phase	Study
31.6%	31.6%	30.8%	Lapworth and Teal, 1994
53%	23%	24%	Goldsmchmidt and Lent, 1997
55.6%	13.3%	30%	Nutting et al.., 1996
68.2%	13.3%	18.5%	Plebani and Carraro, 1997
75%	16%	9%	Stahl et al.., 1998

Table 2.4: Laboratory error frequencies [taken from Bonini et al., 2002].

In conclusion from Table 2.4, most errors occur at the pre-analytic and post-analytic stages. In fact, all of these studies agreed that the majority of errors occur during the pre-analytical phase, but with significant percentages in the other two stages of testing.

Table 2.5 presents some methodological information about these studies, as well as the findings these studies had in terms of impact of errors found on patient outcomes.

Study	Data collection period	Number of samples analyzed	Impact on patient outcomes			
			None	Mild	Moderate	Severe
Lapworth and Teal, 1994	1 year	997000	n/d	n/d	n/d	n/d
Goldschmidt and Lent, 1997	6 years	not determined (n/d)	43%	23%	26%	8%
Nutting et al., 1996	6 months	n/d	13%	13%	n/d	n/d
Plebani and Carraro, 1997	3 months	40490	74%	19.6%	6.4%	0%
Stahl et al., 1998	3 years	676564	n/d	n/d	n/d	n/d

Table 2.5: Patient outcomes in five laboratory error studies (adapted from Bonini et al., 2002).

When considering the impact of laboratory errors found on patient outcome, these studies take account of four levels (none, mild, moderate, severe). These are classified according to immediate impact (i.e. delay of diagnosis) but do not examine the long terms effects of errors identified. This is problematic because it can be hard to determine the degree to which an error affected long term prognoses given that individual patient related factors have an impact on outcomes.

In addition, the Goldschmidt and Lent study [1997] found that approximately 75% of laboratory errors are likely to result to tests which are still within their reference intervals.

This suggests that they would not necessarily have any adverse impact on patient's health, although they might be misleading. However, defining these boundary values is a great challenge, primarily because not all patients are going to respond to the same treatment in the same way.

These challenges limit our understanding of the impact of laboratory error on patient diagnosis and treatment overall. As laboratories maintain little or no information regarding the impact of their work on patients' health, it is very difficult to understand of the severity of laboratory error; a single error may have much greater impact on a patient's health, while several marginal errors may be superficial [Plebani and Carraro, 1997].

One more factor that makes laboratory error difficult to cope with—and, as this thesis argues the most important factor—is that it may be very difficult to detect, not only within the laboratory, but also when these test results are taken into consideration during patient diagnosis within primary care. Erroneous test results which are plausibly acceptable will mislead diagnosis, and therefore the decision upon a course of treatment. In the case of screening, such an error can become detrimental either by missing ill patients or by aggravating their health by e.g., excessive radiation treatment. The next section will therefore discuss the aspect of 'plausibility' of erroneous test results.

2.3.2 Plausibility of false test results

Errors taking place during the testing cycle may not always produce detectable abnormal results, nor raise questions for the physician that has requested them; they can thus be

taken into account during patient diagnosis and decision upon treatment [Bonini et al., 2002]. When false test results appear to be worthy of belief, these are referred to as 'false yet plausible'.

In the case of screening, such errors can take the form of false positives/false negatives:

- A *false positive* is when there is no disease (or other condition) but the results come back as positive; for example, a positive test for HIV or cancer, when the person was disease free, would be a false positive. Sometimes, when a disease is very rare, and/or when a test has a high rate of error, there may be more false positives than actual positives.
- A *false negative* is when there actually is a disease (or other condition) but the results come back as negative [De Boer et al., 2002].

False-positives and false-negatives are a well known problem in laboratory medicine and to a certain level, they are inevitable in any screening programme [Johnson and Patnick, 2000]. In the case of screening, there are two possible outcomes of the analysis: positive, or not positive; both of which are plausible, especially if there is no other information to constitute them as implausible. For instance, when patients are asymptomatic but ill, a false negative will seem like a plausible result.

The incidents that were discussed in Table 1.1 are adverse events that involved several erroneous yet plausibly acceptable test results which were caused by technical faults or systematic errors taking place in the laboratory during highly automated routine analytical processes. Therefore the risks associated with plausibility increase when taking

into account the potential of hiding the automated production of multiple test results, and not just one single erroneous report.

2.4 A pattern of failure in screening services

The result of any screening service will be in the form of a 'positive' or 'negative'. So despite the variety of screening services available, errors in the testing cycle are likely to result to false-positives or false-negatives, which, in most cases seem plausible in isolation for any screening test. Without interaction with the patient or a substantial amount patient history, any deterministic derivation of a screening test will therefore be very likely to manifest in a way plausible for the diagnostician.

In the incidents summarized in Table 1.1, such plausibly acceptable yet erroneous screening tests were being produced for periods of months, or even years, affecting hundreds of patients. While their plausible nature is an inhibiting factor in terms of *detection*, there are some organizational and technological elements that contributed to the *automation* of multiple mistaken tests. In particular, the dependence of several points of care on one single laboratory increases the likelihood of propagation of a single fault to more patients; this becomes more critical when faults in software that participates in the analysis and communication result in the automated creation of multiple false test results.

This chapter discussed how screening is done in the UK, within the organizational structure of diagnostic networks; where several GPs and even other hospitals depend on a single laboratory or radiology department for the screening of all of their patients (Figure 2.2). Such a structure may be useful in terms of productivity and efficiency; however, if

there is a fault in the highly automated testing process (either software bugs or procedural problems), it is very likely that it will affect many—if not all—of the points of care that depend on that laboratory.

The UK incidents of Table 1.1 all developed following this pattern. Moreover, they were prolonged by a lack of (or poor) safeguards—the quality assurance systems in place (primarily audits and management meetings) failed to mitigate or control these errors. As we shall see later on, these incidents were eventually *detected* and *recovered* because people involved took initiatives which were not necessarily prescribed to them.

As mentioned in Section 2.2.2, the screening services quality assurance should be able to "…*deal with errors and adverse events quickly*…" [Programme Director's report, 2005]. However, the fact that several such incidents have occurred since 2000[6] (Table 1.1) indicates that there are several problems that impede timely detection and efficient incident response. This thesis argues that there is a common pattern of failure that may occur in any screening service in the UK. This pattern of failure is summarized in Table 2.6: .

[6] The most recent event took place in 2006 [www.guardian.co.uk/society/2006/jun/29/cancercare.health, last accessed 01-Oct-08]

Fault in the analytic process	Software bugs, human error during analysis, procedural errors in the laboratory.
Automation and propagation	Errors mentioned above are propagated across diagnostic networks.
Plausibility of false test results	False test results are plausibly acceptable, and are thus used in patient diagnosis.
Lack of (or poor) safeguards	Incident reporting scheme not used, audits not carried out—even if they had been requested
Poor incident response	Inability to detect. Also, during the crisis, severe communication breakdowns prolong the incidents.

Table 2.6: A pattern of failure in NHS screening services.

The purpose of this thesis is therefore to understand in detail what underlying problems in the NHS—and screening services in particular—can result to the development of multiple misdiagnoses over prolonged periods of time. The focus then is on the inability to detect and recover from such errors when they occur, but not on the causal factors that led to the errors in the first place. This is because the causal factors vary greatly in the different types of laboratory medicine, phase of testing and types of error that can occur (laboratory error types were discussed in Section 2.3.1). Despite the variety in causal factors, plausibility of false test results and problems in the incident response were somewhat similar. Therefore, the scope of this thesis is limited to the events that take place once false test results begin to be produced and detection is possible.

The continuous effort to remove such *faults in the analytic process* (Table 2.6:) does not necessarily ensure that errors will not occur again the future. With the potential for false yet plausible test results being present, it is essential that diagnostic networks are prepared to detect and recover from such failures as quickly as possible. This is the reason why this thesis focuses on error management within this setting.

The next section will summarize this chapter, before we proceed to the next chapter which will present and discuss the theory that this thesis draws upon: error detection and error recovery.

2.5 Chapter summary

This chapter presented the field and the focus of this thesis. This section will summarize the various concepts and practices that have been presented:

- *Field of study:* The field of this study is diagnostic services (laboratory medicine and radiology) in the NHS, and in particular screening services. This chapter presented the testing cycle which is made up of three stages (pre-analytic, analytic and post-analytic). Then, screening programmes in the NHS, their organizational structure and the role of information technology were discussed. Finally, NHS laboratory quality assurance practices were briefly presented.
- *Focus of study:* The focus of this research is laboratory error, and most specifically the potential for plausibly acceptable false test results. This aspect of laboratory error is crucial because it makes it very difficult to detect errors once they have occurred, so false test results may be used in the process of diagnosis, and consequently have adverse affects on patients' health.

This chapter concluded with a pattern of failure that this thesis has observed in several NHS screening incidents: Multiple, false yet plausibly acceptable, screening tests which were not detected for long periods of time.

Chapter 3: Theoretical context

The previous chapter presented the context of this work, and focused on a pattern of failure that has materialized in several screening incidents, resulting in prolonged failures to detect and recover from laboratory errors. This chapter will present the theoretical background that this thesis draws upon. The literature review will introduce concepts regarding *error detection* and *recovery* that come from cognitive psychology and safety science, which will then be related to laboratory error and its manifestation in test results. This review will conclude with an 'Error Recovery Framework' that will be used throughout this thesis, both for the analysis of the four case studies and the generation of recommendations for the improvement of laboratory error detection and recovery.

3.1 Error recovery versus error prevention

The study of error within cognitive science and applied psychology has made a significant contribution in understanding what types of error may occur, and what types of causes may result to different forms of error [e.g. mistakes, slips and lapses by Reason, 1990]. Within safety research in industry, such work has been very influential. More recently, the study of major disasters (e.g., Chernobyl [Watt Committee, 1988], the Challenger [Vaughan, 1996]) has extended our understanding of error, not only as a human action, but also as a result of a wider, system failure. Error does not only refer to an act that will immediately result in a hazard (e.g., pressing the wrong button), but may

also build up through time through complex organizational and socio-technical environments (e.g., maintenance error) [Reason, 2004].

Safety largely depends on producing systems without defects, while reducing the potential for human errors to occur; this can be seen as 'error prevention', and can be seen a pro-active approach to safety. However, error prevention is not the only strategy towards achieving an acceptable level of system safety [Lewis and Norman, 1986; Frese, 1991]. As it is extremely difficult to remove all potential for errors or technical faults [Greenwell et al., 2004], industrial systems have complex sensor-alarm systems to detect potential problems before they compromise safety. In addition, they employ incident response and crisis management procedures, in order to control failures as they occur and prevent or minimize their impact and consequences. Error recovery is therefore a second strategy (reactive approach) which, along with error prevention, can provide adequate system defences for the prevention and control of accidents [Kontogiannis, 1997].

The two error resistance strategies, along with the concepts of forward and backward error recovery which will be discussed in the following sections are summarized in Figure 3.1.

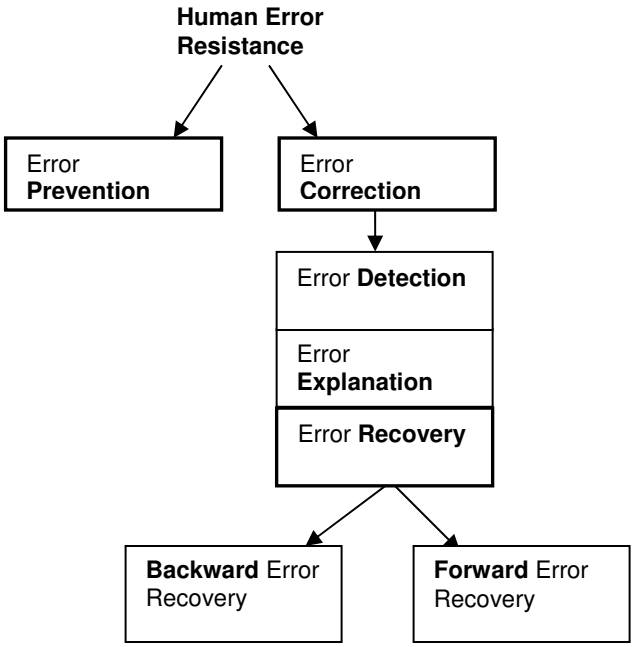


Figure 3.1: Error resistance strategies [adapted from Jambon, 1997].

Error recovery (or error handling) has received less attention in research than the causes human error [Zapf and Reason, 1994; Kontogiannis, 1997]. In a similar manner, there is much more literature available in error prevention than error handling [Klein et al., 2005; Blavier et al., 2005]. Nonetheless, in the past 15 years, there has been some important work in the areas of error detection and recovery (however very limited in the area of patient safety). The purpose of this chapter is to therefore present some of this work, and place it in the context of laboratory screening error.

3.2 Overview of error handling

Improving the ability of a system to detect and recover from errors requires design and engineering effort, as well as organizational and structural considerations. Some system features that take part in the detection and recovery of errors and technical faults include the computer user interface [Rizzo et al., 1996], checking mechanisms [Clarke, 2005],

incident reporting, communication [Barenfanger et al., 2004], work design [Zapf et al., 1994] and training [Chmiel and Wall, 1994].

Research and practice in error handing can be seen to fall under two main categories. These aspects of error handling are the following:

- **Error detection:** This area of research considers the mechanisms through which errors and technical faults are either identified, or merely steer suspicion, and is a focus of cognitive science (Section 3.3).
- **Error recovery:** Recovery follows directly after detection. A variety of models and frameworks describe the stages of activities that make up the process from detection to recovery, coming from systems engineering (Section 3.4).

3.3 Error detection

Error detection is the first step of an error recovery process [Reason, 1994; Kontogiannis, 1997; Klein et al., 2005]. Error detection is a human, cognitive activity, which can be defined as "*the realisation that an error has occurred without necessarily understanding the nature and cause of the error"* [Zapf and Reason, 1994].

Sellen [1994] proposed a theoretical taxonomy of detection modes which aims broadly to describe the ways through which people detect errors in a wide variety of everyday tasks. The framework describes a range of detection mechanisms falling into three categories: action-based, outcome-based or through limiting functions. The three modes of detection, along with an example of laboratory error detection are given below:

- **Action-based detection**: Occurs as the error takes place—this can be best described as "caught-in-the-act". Action-based detection occurs when there is a mismatch between an action plan and the executed actions, or if there is a mismatch between conscious intentions and executed actions. For instance, a laboratory technician detects his or her own error during a specimen analysis by realizing the analysis is not conducted as prescribed. Action-based detection can also occur when a person observes someone else while performing an erroneous action.
- **Outcome-based detection**: This is based on the evaluation of the outcome of the erroneous action. This kind of detection can occur if there is a mismatch between expected outcomes and actual outcomes, or if there is a match between expected error forms and outcome. In a laboratory setting, this can take place when a physician examines a test report and the results do not make sense.
- **Detection through external limiting functions**: Limiting functions refers to physical constraints imposed by the environment. This kind of detection is somewhat different from action-based and outcome-based detection as no evaluation of the correctness of an action is required. In a laboratory context, an example of detection through limiting functions is the following: A ward nurse attempts to phone the pathology department in order to request a test, but he or she has actually phoned the haematology department. This is realized when the person who picks up the phone announces "*you have reached the haematology department*".

In complex environments, error detection may not always occur instantly, but may in fact be a result of several events; even though error detection is nothing more than an acknowledgement of the presence of an error, without knowing what caused it, or even what the nature of the error is. Before being sure that something *is actually* wrong, people may first be concerned, or suspicious that something *might be* wrong. This has been referred to in literature as error suspicion [Allwood, 1984], problem recognition [Cowan, 1986], problem discovery [Woods et al., 1987] and problem detection [Smith, 1989]. The term 'problem detection' will used hereafter, and will be discussed in the following section.

3.3.1 Problem detection

Problem detection is the process by which people first become concerned that events may be taking an unexpected and undesirable direction [Smith, 1989]. Problem detection can occur even in the absence of a fault or the occurrence of an error; however, it signifies the existence of a *potential* fault/error. This is important as in many cases, problem detection may lead to an early resolution of a problem before it manifests into a dangerous condition. Even in a steady state condition, problem detection signifies a preparedness and alertness in case something does go wrong.

Problem detection is a sense-making activity. A person who is concerned may act in many different ways to determine if there is a problem, and what that is. For instance, if problem detection occurs, one might decide to monitor system behavior in case another cue of the potential problem emerges, or might decide to take action despite being sure

that something is wrong. In some cases, an operator may remain suspicious even when concerns are explained away by others [Klein et al., 2005].

The 'discrepancy accumulation model' [Cowan, 1986] is one the first influential models of problem detection, even through there had been some earlier propositions [e.g., Davies 1973]. Cowan described problem detection as "*the accumulation of discrepancies until a threshold was reached*". In a laboratory setting, the accumulation of discrepancies can for instance be seen as a growing concern over the frequency of positive or negative test results, without however being sure that an error has occurred.

Klein et al. [2005] have suggested problem detection is affected by the following three factors:

- **Expertise:** Expertise can be an advantage in most cases. Skilled personnel are most likely to generate expectancies and be decisive to take action when contradictions occur. A skilled operator is most likely to have an understanding of what conditions can result in the system generating misleading readings. On the other hand, expertise can result in confidently explaining away a problem. For instance, an experienced nurse may become concerned when he or she notices a discrepancy in the frequency of positives/negatives, while an experienced laboratory technician might not find this alarming.

- **Stance:** Stance refers to the orientation the person has towards the situation [Chow et al., 2000]. Stance can be an absolute denial that anything could possibly

be wrong, to a hysterical over-reacting attitude towards minor signs. General alertness, level of suspicion and emotional status are some of the factors that make up a person's stance. For example, a nurse that is hesitant to trust new technologies in his or her workplace will be 'on the lookout' for any problems a new system may incur. In healthcare systems, information is evaluated upon the basis of the "perceived credibility of the source" [Cicourel, 1990]. In other words, the judgment over the credibility of a test report will be made by taking into account the level of trust in the laboratory department that produced it. If there is no prior experience of problems originating from the source of information, it is unlikely that the plausibility of a test result will be questioned.

- **Attention Management:** This refers to sensor/alarm systems: having system facilities to detect, capture and notify operators about problems. Attention management is however not independent of expertise and stance, as they may for instance result in disregarding an alarm or remaining concerned in the absence of an expected alarm [Wickens and McCarley, 2007].

It is very likely that once a specific error has occurred, problem detection will occur before error detection—although this is not always necessary. In most environments, operators that have detected an error are expected to report this or take immediate action if it is within their responsibilities. However, the uncertainty that characterizes problem detection can prevent an operator from taking action. In the case that concerns are reported to someone else, it is possible they will be disregarded in the absence of

convincing evidence. For this reason, following problem detection, operators may decide to continue monitoring system behavior for further cues, or may formulate and explore hypotheses as to what is wrong themselves [Klein et al., 2005].

In everyday life problem detection occurs frequently, and, in many cases, instinctively. It could be argued that when, for instance, driving a car, a driver's decision to slow down is dictated by problem detection as, not slowing down will result in a car crush. In a similar context, if while driving a car the driver hears a "strange" noise coming from the engine, it is possible that he or she will not take the car to the mechanic until that sound occurs again.

The following section will present a number of models and frameworks that describe the process of error recovery; the process that follows error detection (but not problem detection) until recovery (or giving up).

3.4 Error recovery

Error recovery is generally made up of three stages: Error detection, error diagnosis, and error correction [Bagnara and Rizzo, 1989; Zapf and Reason, 1994; Jambon, 1997]. This generic process can be applied to any type of error within almost any context. However, depending on the environment, error explanation and error correction may be broken up further (for instance, the involvement of multiple people in an error handling process will require communication (e.g. incident reporting) or planning might have to take place before proceeding to corrective actions).

3.4.1 Forward and backward error recovery

A distinction can be made between forward and backward error recovery [Dix et al., 1993]. This distinction—used extensively in interactive systems design—refers to the path that can be taken towards recovery [Jambon, 1997]:

- **Backward Recovery:** Backward recovery refers to attempt to restore the system state following the occurrence of an error by following events as they occurred in reverse. According to Yang [1992], there are three kinds of backward error recovery: *undo*, *cancel* and *stop*.
- **Forward recovery:** During forward recovery the operator has to perform unexpected tasks to recover the fault, perhaps through improvisation. This kind of recovery applies mostly to failures in industrial engineered systems, where undo and cancel cannot easily be implemented.

3.4.2 Error recovery frameworks and models

In this section we will discuss the most influential theoretical frameworks that describe the process of error recovery. There are several frameworks proposed since the 1980s , (e.g. [Cowan, 1986]) — the ones discussed here are some of the more recent ones that are seen as a step further from the first attempts that were made three decades ago. In addition, the three frameworks discussed in this section take slightly different perspectives which are of interest to this thesis.

In the following pages, some frameworks will be presented; a discussion regarding their benefits and limitations will then follow in a separate section, which will be based on a

comparison of their features and applicability on the particular matter of laboratory error handling.

Error handling model

One of the earliest models proposed was by Zapf and Reason [1994]. This model (Figure 3.2) views error recovery as a two-step process: error diagnosis and error recovery. Each of these stages is broken down to two further steps.

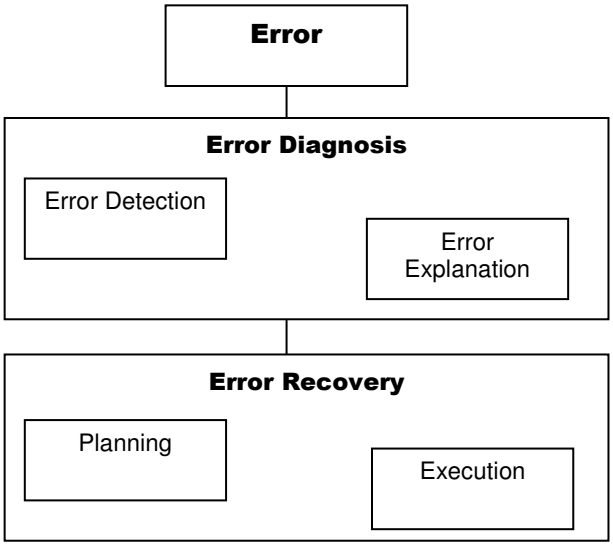


Figure 3.2: A generic error recovery model [adapted from Zapf and Reason, 1994].

Here, error detection is seen as part of error diagnosis, which is then followed by error explanation. Following error diagnosis, error recovery is performed in two steps: planning and execution. The two authors do not include the achievement of error control and correction in their model, as they view that as the outcome of the actual error recovery process (regardless of whether it has been successful or not), taking a rather general perspective on the entire process.

The INCORECT framework

INCORECT (Investigating Cognitive and Recovery Tasks) has been proposed for the analysis of cognitive reliability in interaction in complex systems [Kontogiannis, 1997]. The framework takes into account a set of taxonomies of cognitive error modes, error causes, problem solving failures, recovery mechanisms and contextual factors.

INCORECT has been proposed in order to identify cognitive errors at the stages of interpretation, decision making and planning during the resolution of an encountered problem. The framework can thus be used in a risk assessment context, attempting to identify potential pitfalls such as an unsafe intervention that may result from a wrong situation diagnosis.

Apart from the investigation of cognitive errors, INCORECT also examines the activities that focus on the control of the error's consequences. As we will see later on, this second aspect of this framework relates greatly to this thesis. Figure 3.3 illustrates two dimensions of the error recovery management: Error handling (during a crisis), and the consideration of adequate interventions to increase preparedness for more efficient error handling in the future.

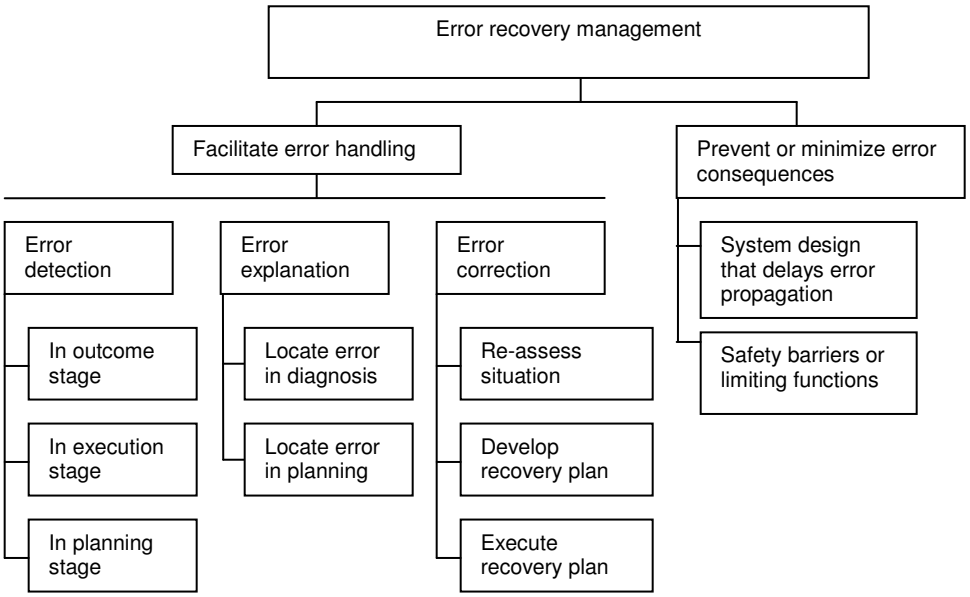


Figure 3.3: The INCORECT framework [adapted from Kontogiannis, 1997]

The human redundancy framework

This framework [Clarke, 2005] is concerned with the role of human redundancy in engineered systems in high-hazard industries. More specifically, the human redundancy framework considers "redundant arrangements" such as operator/supervisor within a socio-technical system where the following features are important:

- Someone checks someone else's work
- A check is carried out at the time a function is fulfilled or soon after it is fulfilled
- The checker is directed, either verbally or through a written procedure, to check a particular human interaction
- The check takes place during normal operation.

According to the author, human redundancy is activated when error recovery commences (Figure 3.4). From an error recovery perspective, human redundancy exists "*where there*

is support for concurrent human recovery by another of an error associated with a required operator function".

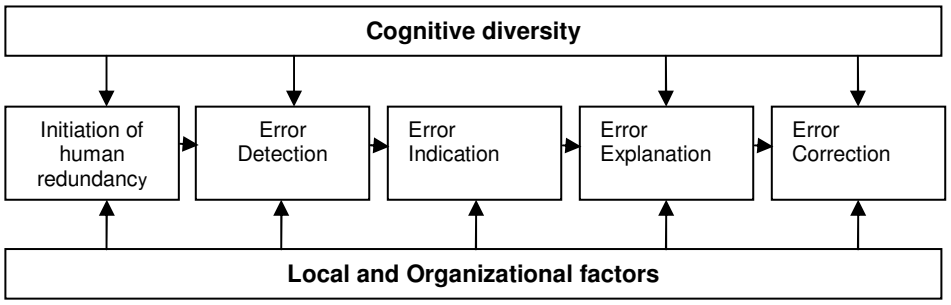


Figure 3.4: The 'human redundancy' framework [adapted from Clarke, 2005].

Each of the stages of error recovery within this framework are affected by cognitive diversity and the local and organizational factors that shape the context within which these activities take place:

- **Cognitive diversity** is the availability of different cognitive behaviors to fulfill a required function, where differences originate either within operators or within their environment. Cognitive diversity may exist with respect to two or more individuals within the same group, between a group and another individual and between two people that perform checking tasks.
- **Local and organizational factors** refer to any element of an organization that may affect error recovery; For example, issues such as stress, over-trust, and deficiencies in resource.

3.4.3 Comparison and criticism

The purpose of this section is to discuss the strengths and weaknesses of each of these three models. This criticism has been based on a comparison of these models and

frameworks, whilst taking into account some of the initial observations regarding laboratory related incidents that have been discussed in the first two chapters.

The error handling model [Zapf and Reason, 1994] can be seen as the most generic one of the three. Although it can be argued that it covers all aspects of the recovery process, its relatively high-level view does not allow for a detailed consideration of all of the activities that may take place during error handling. In contrast, the other two frameworks take more in-depth perspectives:

The INCORECT framework [Kontogiannis, 1997] takes account of various error detection types (similar to the ones proposed by Sellen, 1994), but also includes a more elaborate breakdown of error explanation and error recovery—in addition, this model extends to include error/disaster management activities following the incident (which will be the focus of the next section). One limitation of INCORECT is that it does not take account of communication activities; communication breakdowns are an important aspect of error handling failure and are part of the thesis focus.

However, communication events are part of the human redundancy framework. Error indication—a sort of incident reporting—is essential in such a model, especially if the context is a complex organizational setting, where the involvement of several departments, or even organizations may be possible. Communication during error handling can be crucial, especially in the case of a communication breakdown which could have serious consequences on a recovery process.

The human redundancy framework [Clarke, 2005] considers error recovery as a human activity, whereby detection by others and reporting of errors to persons who can take action are the underlying concepts. This framework also considers cognitive diversity and organizational factors as determinants of the outcome of an error recovery process. However, this model is somewhat limited in the variety of possible types of error detection, and error explanation, which are explained in more detail in INCORECT.

A common issue in these models is that they do not consider problem detection as part of the error recovery process. Arguably, a recovery process may commence once the presence of an error has been acknowledged; however, in many cases an error might not be obvious enough for people to confirm its presence; and although the importance of problem detection has been recognized, current recovery models do not incorporate problem detection. Finally, these models do not include the investigation activities that may take place at any time (either to find out if there is an error or not, or to identify the causes of the error).

With this in mind, it was found necessary to proceed with a new model that considers problem detection and further investigation as parts of the error recovery process formally. The following section will present this new model which builds upon INCORECT and the human redundancy framework, also taking into account problem detection.

3.5 The error recovery framework

Figure 3.5 illustrates the stages that make up the error recovery framework. This framework indicates the sequence of events from problem detection to error correction; it includes a larger set of steps to be followed, providing more coverage in terms of types of error handling activities that are possible.

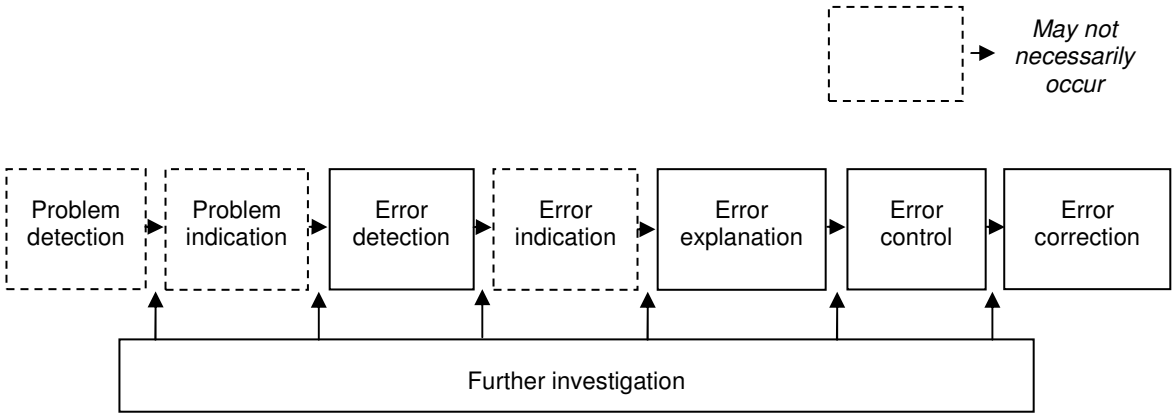


Figure 3.5: The error recovery framework

These stages can be summarized as follows:

1. **Problem detection:** The initial concerns that events may be taking an unexpected and undesirable direction that potentially requires action.
2. **Problem indication:** The reporting of concerns to someone who can act upon this potential problem.
3. **Error detection:** The realization that something is actually wrong, without necessarily knowing the nature of the error or what has caused it (Action-based, outcome-based, through limiting functions).
4. **Error indication:** The reporting of error detection to someone who can act upon the error.

5. **Error explanation:** The localization of an error and the identification of its causes.
6. **Error control:** The controlling of the incident by stopping the error from causing more harm. At this point, the error has not necessarily been removed from the system.
7. **Error correction:** The causes of the error are being removed and system can operate normally again.
8. **Further investigation:** This can take several forms (e.g. observation of system behaviour, enquiry to colleagues, examination of system features, technical analysis, etc.).

Problem detection, problem indication and error indication may not always occur; however these activities would take place in situations where there are multiple people involved (in terms of indication) and where system feedback is limited, making error detection more difficult. The error recovery framework also takes into account Sellen's [1994] taxonomy of error detection mechanisms. Although cognitive diversity is obviously very likely to determine the outcome of each of these stages, it is not considered in this framework. This is because it is very difficult to map the cognitive model that each involved person would have during error recovery without the appropriate amount of data.

The error recovery framework will be used throughout this thesis in order to describe error handling activities in the context of laboratory error handling. The way with which

it will be used will be discussed in more detail in the next chapter (Chapter 4: Research methodology). However, before proceeding to the next chapter, the next section will discuss how national and corporate regulation and policy determine how organizational failures are managed, and how organizational learning towards better error handling may be guided by formal public inquiries in the UK.

3.6 Chapter summary

This chapter presented and discussed existing work in the areas of problem detection, error detection and error recovery, which form the theoretical basis that this thesis draws upon. Following the presentation of various models and frameworks, this chapter proceeded with their criticism which concluded with a new error recovery framework—an adaptation of existing frameworks which also takes into account problem detection as part of error recovery.

Chapter 4: Research methodology

The previous chapter presented an overview of existing literature in error detection and error handling, and concluded with a proposed error recovery framework which will be used to analyze and explain the activities that make up the processes and activities of laboratory error handling. This framework will be the basis of the research methodology that this thesis undertakes. The methodology is summarized in Section 4.1.

Section 4.2 presents an overview of available techniques that were considered for this analysis, while Section 4.3 contains a discussion based on criticism of these techniques. In order to achieve the objectives of this thesis, a new technique has been developed, by adapting STEP [Henrick and Benner, 1983] and integrating it with the error recovery framework that was proposed in the previous chapter.

Finally, Section 4.4 will discuss how this suggested analytical approach will be applied to multiple case studies, and Section 4.5 will discuss the sources of data that these analyses have been based on.

4.1 Overview of research methodology

The research methodology that this thesis follows is presented in Figure 4.1, which also illustrates the relationship between methodology, research objectives and overall thesis aims.

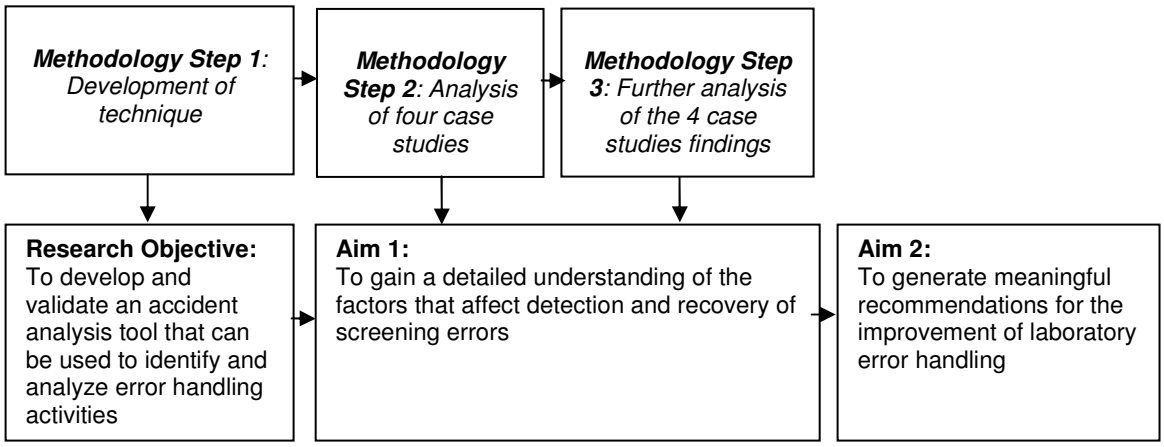


Figure 4.1: Overview of research methodology, aims and objectives.

The research methodology is therefore summarized as follows:

- **Step 1:** *Development of technique.* A technique is developed and discussed in this chapter. This technique is a result of the integration of the error recovery framework that was discussed in the previous chapter (see Section 3.5) with an existing accident analysis technique.
- **Step 2:** *Analysis of case studies.* The UK incidents summarized in Table 1.1, along with one USA incident will be analyzed with the technique that has been proposed. This incident was included as a first attempt to demonstrate the generic nature of the approach and was conducted as a result of cooperation with a major North American hospital department.
- **Step 3:** *Further analysis of the four case studies findings.* The findings of the four case studies will be further analyzed in order to identify key problem areas

(common problems in these case studies) and draw high-level conclusions about laboratory screening error handling. This will result in an enriched model of the error recovery framework which will describe the different types of error handling activities that may take place within a laboratory screening context.

4.2 Accident analysis

The aim of accident analysis is to identify the underlying causal factors that shaped the events of the misfortunate outcome. Such analyses can in turn be used by technology designers, system engineers, management and regulatory authorities to consider how to eliminate these underlying factors so that similar occurrences are avoided in the future. With this perspective, investigations into laboratory failures can be used by the laboratory and hospital management, and by device manufacturers for the continuous improvement of their respective products [Jenny and Jackson-Tarentino, 2000].

An abundance of techniques and notations have been developed by authorities and researchers, considering various levels and aspects of different kinds of incidents and accidents in safety-critical industries[7]. Causal analysis can be used to explain *why* the failure took place, and reconstruction techniques can explain *what* happened during the failure. Other approaches, such as Management Oversight and Risk Tree (MORT) [Johnson, 1973] and Systems Theory Accident Modelling and Process (STAMP) [Leveson et al., 2003] examine the involvement of organizational, human and technological elements in the occurrence of an incident or accident.

[7]This chapter does not present a review of all accident techniques, as such is already available. The "Handbook of incident and accident reporting" [Johnson, 2002] has been used to guide the discussion that takes place here. However, event-based techniques will be considered in more detail.

However, despite the fact that accident analysis techniques would identify and analyze causes and events that resulted in poor error handling, there is no systematic approach taking a focus on the organizational response following error detection. It is argued here that having a focus on error handling in accident analysis would help us to understand how the error could have been detected faster and recovered, in addition to the why and what happened that established accident analysis notations can help identify.

Error recovery is a process evolving through different stages over time [Zapf and Reason, 1994; Jambon, 1997]. For this reason, it was more appropriate to consider event reconstruction approaches for the development of a recovery-focused analytical tool. The following section will present and discuss some event-based techniques which have been considered for this analysis.

4.2.1 Event-based techniques

Event-based techniques are used to model multiple events which are linked over time. Most event-based techniques are supported by a graphical notation which depicts the evolution of events in a left-to-right, linear manner. The first event in a chain is often referred to as the "initiating event"—there is however no principle that dictates the selection of the initiating event [Leveson, 2001]; this decision is largely subjective—the investigator may go back several years before the occurrence of the more immediate events that led to an accident/incident.

Event-based techniques are useful in understanding the mechanism through which an accident/incident was created; however a common criticism of such techniques is the limited insight they provide in analyzing the underlying causes and their conditions [Johnson, 2003]. For the reason, event-based models are used to guide further analyses of the particular events that were identified as determinants in the sequence of the adverse events.

This section will present and discuss the following:

- Events and Causal Factors Analysis (ECFA), [Buys and Clark, 1995]
- Multi-linear Events Sequencing (MES), [Rimson and Benner, 1975]
- Sequentially Timed and Events Plotting (STEP), [Henrick and Benner, 1983]

There are several event-based techniques available. The particular techniques have been selected because they are well established, longstanding and well documented.

Events and Causal Factors Analysis

Events and Causal Factors Analysis (ECFA) is used to identify errors, changes, oversights, and omissions, as well as also the relevant conditions affecting each event in the accident sequence. The approach breaks down the sequence into a logical flow of events from the beginning of accident development. The end point may be defined either as the loss event itself or as the end of the amelioration and rehabilitation phase. In addition, this flow of events may not necessarily lie in a single event chain but can involve several confluent and branching chains.

Figure 4.2 illustrates the conventions that are used when applying ECFA.

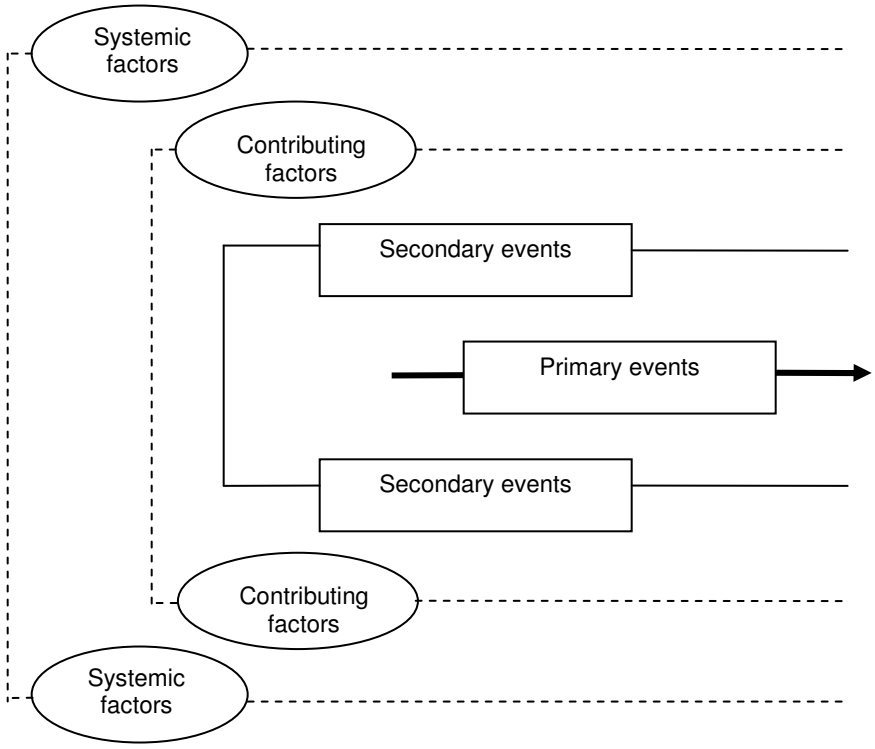


Figure 4.2: ECFA notation [adapted from Buys and Clark, 1995].

ECFA is designed as a stand alone technique but is usually applied with other techniques found in the Management Oversight and Risk Tree (MORT) programme [Buys and Clark, 1995]. ECFA serves three main purposes in investigations:

- assists the verification of causal chains and event sequences
- provides a structure for integrating investigation findings
- assists communication both during and on completion of the investigation.

Multi-linear Events Sequencing

Multi-linear Events Sequencing (MES) is a method that is made up of concepts, principles, rules and procedures which can be used for any kind of investigation. The technique was developed by Rimson and Benner—at the time investigators with the

National Transportation Safety Board (NTSB) [1975]. One of the motivations behind the development of MES as an event-based technique was the wish to avoid the use of checklists, as this may allow for factors which are not included in a checklist to be missed [Ferry, 1998].

MES uses a matrix-based structure which consists of data documentation, organization and analysis tools and rules in order to drive the investigation tasks. Matrix entries follow pre-defined grammar and syntax rules of construction for event blocks on matrices including person, number, tense, voice and deictic position—the MES data language—and reasoning rules to develop tested descriptions and explanations of what happened. These descriptions are then analyzed systematically with orderly sequential problem defining. MES provides some generalized behavioral models, guiding principles and assessment or ranking tools to convey knowledge from prior experiences to help investigators [Benner, 2003].

Figure 4.3 illustrates an example MES chart.

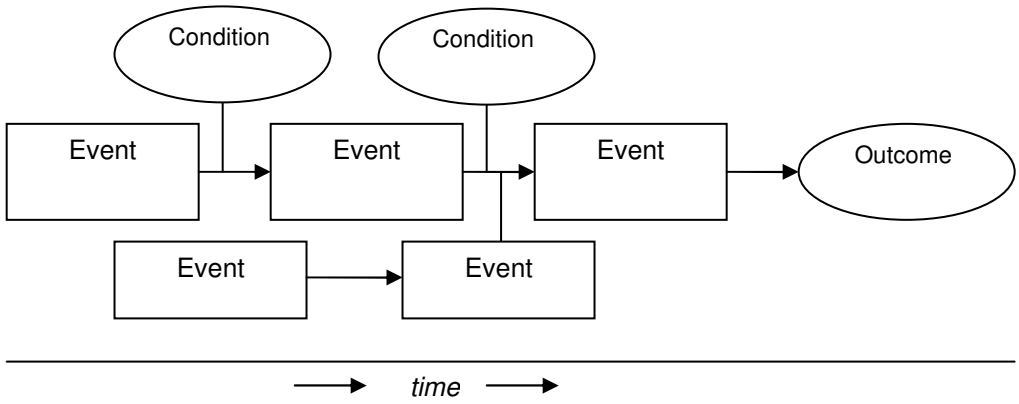


Figure 4.3: An example MES diagram [adapted from Ferry, 1998].

Every event is a single action by a single actor. The actor is something that brings about events, while actions are acts performed by the actor. A time line is displayed at the bottom of the chart to show the timing sequence of the events while conditions that influence the events are inserted in the time flow in logical order to show the flow relationship.

Sequentially timed and events plotting

Sequentially timed and events plotting (STEP) [Henrick and Benner, 1983] can be seen as synthesis of ECFA and MES [Johnson, 2003]. The starting point of STEP is the compilation of STEP cards—cards that provide an initial means of recording information about key events that occur during the course of an incident (for example, see Figure 4.4 below).

Event card id:	
Actor:	
Action:	
Event location:	
Time/date event began:	
Event duration:	
Data source:	
Description:	

Figure 4.4: Example STEP event card identifier [adapted from Henrick and Benner, 1983].

One of the criticisms of ECFA and MES is that it can lead to very complex charts, which are difficult to maintain without tool support. The use of STEP cards aims at minimizing the notational excesses of the other analytical techniques. The multi-linear time-based event representation is conducted in a similar manner as MES, although somewhat more simplified (with condition events now omitted). Figure 4.5 illustrates an example STEP diagram.

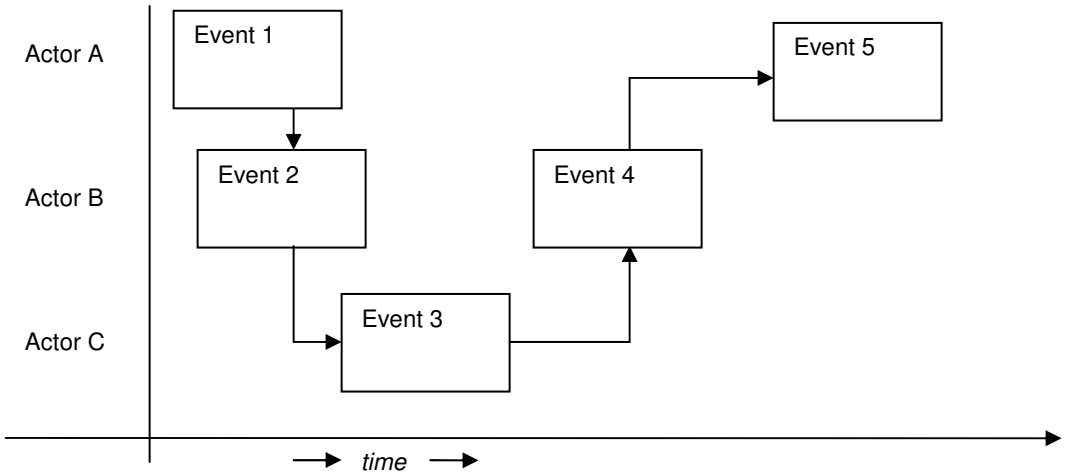


Figure 4.5: Example STEP diagram.

4.2.2 Focusing accident analysis on error recovery

Event-based techniques can be used to understand the mechanisms that, through time, resulted in an accident or incident [Henrick and Benner, 1983]. The time-based, graphical illustration of events can help the analyst reconstruct what happened, and further examine why these events occurred. The techniques discussed in this section (ECFA, MES and STEP) are long-recognized, established, event-based analytical techniques. However, these techniques do not distinguish between the events that led to an error, and the events that followed aiming at preventing or controlling the impact of the error once it has occurred.

Error recovery is itself a process that evolves through time [Zapf and Reason, 1994; Kontogiannis, 1997]. Therefore, an event-based technique could be adapted to focus on error handling, and the development of an entirely new technique was not found to be necessary. In order to focus on error handling activities, one has to filter out these events and isolate them from other events. This can be done by the use of the error recovery

framework (or any other error recovery framework) which was proposed in Chapter 3. Essentially, the error recovery framework may be used in combination with any event-based technique. However, for the purposes of this analysis, STEP was found to be more practical, due to its simplified notation, and the use of STEP cards.

The following section will present the result of the integration of STEP with the error recovery framework; a method which has been developed for the purposes of this thesis, and will be used throughout the analysis of the four case studies.

4.3 Error recovery focused STEP

As discussed previously, one of the primary criticisms of ECFA and MES is that their notational complexity limits the ability to effectively manage extended diagrams without the support of software tools. The decision to use the error recovery framework within an event-based technique therefore required a simpler technique that could then be enriched with various recovery event-type definitions. The ability to separately document events and information about them with STEP cards allows for an event-centric elaboration without overloading the graphical event chain.

The integration of STEP with the error recovery framework (Section 3.5) resulted in Error recovery focused STEP[8] (ER-STEP). Analysis of error recovery activities with this technique is performed in the following stages:

1) **STEP event cards** are produced as in the STEP method.

[8] The technique and its application resulted in the following paper: Chozos, N (2008). Focusing accident analysis on error handling activities: Three case studies in the NHS. Reliability and Engineering, Special issue [accepted for publication]

2) **Error recovery events** are identified according to the definitions provided in Section 3.5: the error recovery framework. In the action cell of the STEP cards, the recovery stage name may be entered. For error detection events, the action-based, outcome-based, through limiting functions definitions can also be used for more detail. Events that are not recovery related may be labelled as "no classification".

3) **ER-STEP diagrams** are drawn up in a similar way as STEP diagrams. The error recovery stage name should also be entered in the event boxes. It is suggested that different colours are used for each of the stage, as this can help to visually observe the process and easily draw conclusions (e.g. the frequent occurrence of a specific type of event) but these could be omitted. Figure 4.6 illustrates a suggested coloration of the error recovery framework stages, which will be used in the rest of this document.

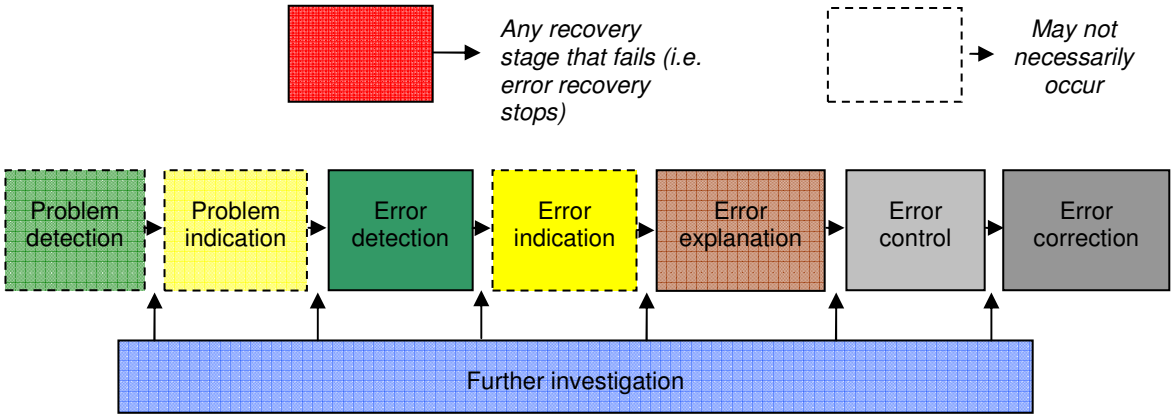


Figure 4.6: The ER-STEP notation (colors used to denote recovery stages).

The "initiating event" is the first instance of detection (problem detection or error detection). The investigator may chose to have separate ER-STEP diagrams for each sequence of events that is initiated by a detection event, but that is not

necessary as several detection events may take place within the same error recovery process. This decision may therefore be made arbitrarily.

4) **ER-STEP diagrams are rearranged** with a single, linear, error recovery view. During a recovery, it is very likely that there are repetitions of a specific kind of recovery action, or that the process enters a loop for a certain time period. In a time-based sequence, all events are depicted with a left-to-right view regardless of any iterations that may occur; if these events are rearranged with an error-recovery perspective, the investigator can see these iterations as a way of 'going a step back', which indicates that the specific recovery stages may be problematic.

The following section will present a simple example of an imaginary scenario which has been analyzed with ER-STEP.

4.3.1 An example of ER-STEP analysis

Let's consider the following scenario (for brevity, recovery event identification will be done in the scenario text):

Car breakdown incident

"On the morning of December 12th, John was driving his car to work as usual. At approximately 8.34am, he heard a strange noise from the car engine [**Event 1: Problem detection**]. He then heard this noise again when parking his car outside his work place (8.45am) [**Event 2: Problem detection**]. At that point, John was a bit worried, although the car did not appear to have any particular problems as he was driving it. He opened the bonnet to check the engine **[Event 3: Further investigation**] but didn't see any problems [**Failure of further investigation**]. After finishing work, John got in his car and started driving back home (5.00pm) [**Event 4: No classification 2**]. That noise did not occur again that day, or the following day. However, on December 14th (5.30pm) on his way back home from work [**Event 5: No**

classification], he heard that noise again [**Event 6: Problem detection**], and before managing to pull over to the next hard shoulder [**Event 7: No classification**], there was smoke coming out his engine [**Event 8: Error detection**]. The noise had stopped, but he was unable to start the engine afterwards [**Event 9: Error detection**]. He called the breakdown assistance [**Event 10: Error indication]**, which arrived shortly after (5.45pm) [**Event 11: No classification**]. The mechanic checked the engine [**Event 12: Further investigation**] to find out that the head gasket was damaged [**Event 13: Error explanation**], allowing for a cooling failure; however, the temperature indicator had not shown an increase in temperature [**Event 14: error explanation**]; the temperature indicator was damaged as well. John then had the car towed to the garage (6.00pm) [**Event 15: no classification**] which was fixed the next day [**Event 16: error correction**]. John realized he should have had the car checked out at the first instance he heard that noise".

ER-STEP analysis

STEP cards

(For this scenario, the second step—identification of recovery activities—has already been conducted so the error recovery stage tag has already been entered in the action box). Figure 4.7 illustrates two STEP cards as an example.

Event card id: **Event 1**	
Actor:	John
Action:	Problem detection
Event location:	Highway
Time/date event began:	Dec 12th, 8.34am
Event duration:	1 minute
Data source:	
Description:	John hears noise from car engine

Event card id: **Event 3**	
Actor:	John
Action:	Further investigation
Event location:	Highway
Time/date event began:	Dec 12th, approx 8.46am
Event duration:	5 minutes
Data source:	
Description:	John checks car engine

Figure 4.7: Example STEP cards.

ER-STEP diagrams

Figure 4.8 and Figure 4.9 illustrate two ER-STEP diagrams that describe some of the events that took place during the car breakdown incident.

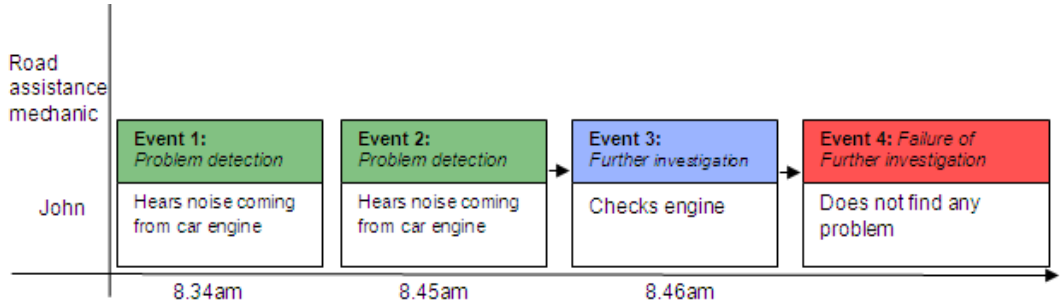


Figure 4.8: Example ER-STEP diagram (1).

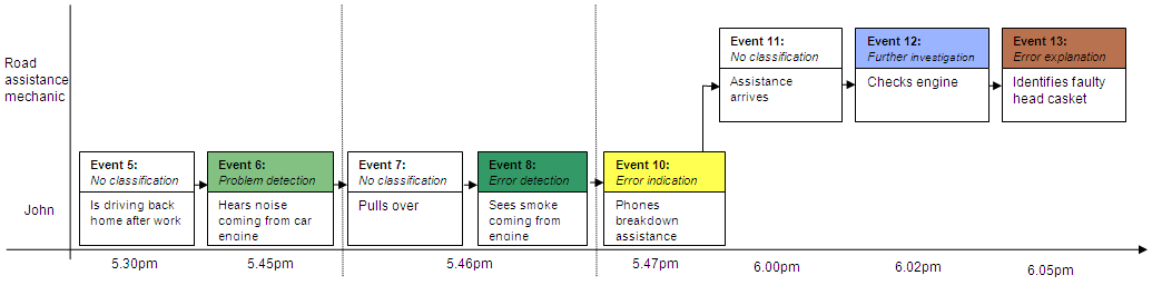


Figure 4.9: Example ER-STEP diagram (2).

Recovery-focused view

Having drawn up the ER-STEP diagrams, the analyst can now isolate error recovery activities even further, and observe the recovery sequence by accumulating and rearranging the recovery events according to the error recovery framework sequence of events (Figure 4.10).

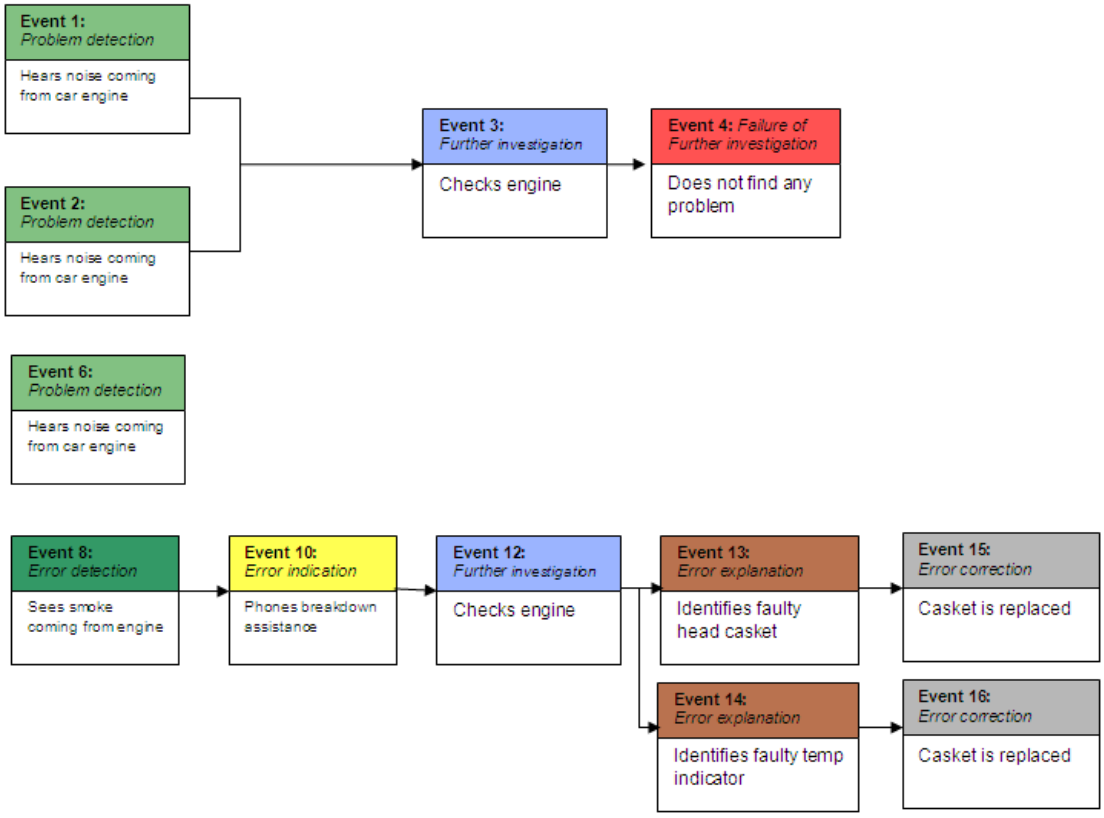


Figure 4.10: Error recovery-focused view.

This view becomes particularly useful when multiple recovery processes have taken place during an incident and when ER-STEP diagrams start becoming large and complex. With this view, discussion per error recovery stage is possible, while some conclusions can be drawn visually; for instance, in the car breakdown example, we can see that problem detection occurred three times but they all were done in the same way when the 'operator' heard a noise coming from the engine. However, this view can be most useful for the comparison of multiple case studies, as they can be viewed in parallel.

4.3.2 Validation of technique

One of the key challenges in successfully applying this technique is getting the recovery event identification right. For this, the definitions of each recovery stage need to be clear

and precise so that the analyst does not confuse events when labelling each one of them with the appropriate recovery stage name.

The definitions that were provided in Section 3.5 were refined during the application of the technique with case studies. However, it was important that other analysts used the technique, and that the results of their application were taken into consideration towards the finalization of the ER-STEP method. To this purpose, an exercise was put together, which was then distributed to four participants. This happened in two stages: First, two accident investigators used the technique and identified some key issues that needed to be addressed. Having considered their comments and updated the technique and the instructions for its application, a second version of the exercise was developed and given to two healthcare professionals.

It should be acknowledged that the number of participants is limited, and therefore the evaluation of the technique may not be sufficient. However, this exercise has been—at least—useful for making some improvement to the technique. In addition, it could be argued that its extensive application with four complex case studies and the quality of the subsequent conclusions and recommendations may stand as validation activities and results. The limited validation is also justified by the difficulty of finding individuals who are experienced in accident investigation and in healthcare – the technique cannot simply be tested on large numbers of undergraduates expecting the same results; although these tests might be conducted in the future – the initial evaluations were used in a formative way to inform the subsequent application of the approach in the rest of the thesis.

The final exercise can be found in Appendix B. A discussion regarding the evaluation of the technique takes place in Chapter 8:.

4.4 Analysis of multiple case studies

As mentioned previously, the four incidents that are summarized in Table 1.1 have all been analyzed with ER-STEP. Using the same technique for the analysis of multiple case studies offers a constructive way to compare and integrate the findings of the individual case studies. In turn, this can help identify key issues that are common to these incidents, and draw high-level conclusions about error handling in screening services in general. These key findings will be discussed in Chapter 6.

4.5 Data collection

The analysis of the three NHS case studies was based upon reports that were produced by the subsequent inquiry committees. All reports have sections which discuss the events that took place with dates and where available the time of the event. Press reports about these incidents were examined, but they did not form part of the analysis.

Apart from the three UK incidents, there is one more case study which describes an incident that took place in Florida, USA. Information about this incident[9] was collected with interviews conducted during a 2-month stay in the University of Florida[10].

[9] Some information regarding this incident can also be found at: http://catless.ncl.ac.uk/Risks/23.64.html

[10] The 2-month stay was possible through the 'Ken Browning Scholarship in Computing and Medicine, 2005'.

4.6 Chapter summary

This chapter presented the research methodology that has been applied in this thesis. An accident analysis tool has been developed, which is used to explore the organizational response to different kinds of error detection. From the abundance of available accident analysis techniques, it was found that none isolate error handling activities and investigating their relationship. Such a focus is necessary in order to analyze error recovery processes and have a consistent way for drawing conclusions about the factors that limited their effectiveness.

Such a technique has been developed in this thesis, and is the central point of the research methodology that is undertaken here. This chapter therefore presented this technique, which is an adaptation of STEP; with a focus on error recovery, the adapted method is called ER-STEP. An example was then used to illustrate how the technique may be applied. Some discussion took place regarding the approach to the technique's validation and evaluation, which will be concluded in the final chapter of this thesis. Finally, the way with which data collection was carried out was also discussed at the last section of this chapter.

The next chapter will present the analysis of four adverse incidents that involved screening errors with ER-STEP.

Chapter 5: Case studies

This chapter will present the detailed analysis of each of the four case studies with the ER-STEP method which was presented in detail in the previous chapter. Each of the four sections of this chapter corresponds to one of the four incidents, and will consist of an incident summary, the detailed ER-STEP analysis and a summary of each analysis. Note that the discussion regarding the factors that affected error recovery in these four incidents is limited here as more discussion will take place in the next chapter. The purpose of this chapter is only to present the detailed analysis of each of these incidents.

The STEP event cards can be found in Appendix A. Although any necessary information should be visible in the ER-STEP graphs throughout this chapter with elaborate explanations of each event in the narrative, STEP cards may at any time be used as reference, especially when considering the 'error recovery-focused view', which does not illustrate the actor, time and location of the event.

5.1 Case study 1: Down's screening errors, Sheffield

This incident was the first one identified and has motivated and guided much of this study. The inquiry report [Ferres et al., 2001] is very detailed and contains a well documented timeline of activities and conversations that took place throughout the five months that the incident lasted. Also, this incident has the greatest involvement of

software and has unveiled some important issues in comparison to the other three incidents.

5.1.1 Incident summary

The first case study describes an incident that took place in Sheffield during the first five months of 2000. The Immunology Department of Northern General Hospital was using a software application (developed in-house 12 years prior to the failure), which, on January 1st, 2000 was affected by the millennium bug. The error was not recovered until May 23rd. Until that time, 158 women had been screened incorrectly for the likelihood of giving birth to children with Down's syndrome. Out of the 158 women, two eventually gave birth to children with Down's, and two proceeded to a late abortion (Table 5.1 presents an overview of the incident).

Incident	Errors in Down's screening for pregnant women
Incident timeframe	January 1 – May 23, 2000
Primary cause	Millennium bug affected software algorithm used in Down's screening
Data source	*Formal inquiry report* [Ferres et al., 2001]

Table 5.1: Overview of incident 1.

The incident was severely prolonged by a poor organizational response which is also the focus of the analysis. The inquiry committee that was subsequently formed to investigate the errors placed much of the focus of the investigation on "*…determining at what stage following 1st January there were indications that there was a serious problem with the Downs Screening program and how such concerns were addressed*".

5.1.2 Background

Down's screening in Sheffield

The normal process of Downs Screening is a complex multidisciplinary process involving obstetricians, radiologists and the diagnostic laboratory. The initial steps are taken at the time of a woman's first visit to the antenatal clinic when, following counseling, consent for several screening procedures is obtained. Usually this is at 12-13 weeks of gestation. Around this time the woman also undergoes ultrasound scanning which has the dual purpose of identifying foetal anomalies and providing a estimate of foetal age based on measurements of foetal size.

Several blood tests are taken at 15-17 weeks which are sent to the laboratory for analysis.

Complex calculations are then used to estimate the risk of the foetus being affected by Downs syndrome. These start with the age-related risk (the *a priori* risk) which is derived from known incidence of pregnancies affected by Downs syndrome based on maternal age. This risk is modified using a likelihood ratio of the presence of an affected foetus derived from the concentrations of the markers to provide the final risk value. The calculation relies critically on accurate estimation of the projected maternal age at delivery which, in turn, is based on the ultrasound measurements of foetal age and the date of birth of the mother. It is these calculations that the software system in question was used to perform.

The final risk value is reported to the obstetrician who will offer to women deemed at high-risk a definitive diagnostic procedure based on cyto-genetic analysis of amniotic fluid obtained at amniocentesis. The risk cut-offs used in most centers for the latter decision lies within the range from 1 in 200 to 1 in 300. This value is selected due to the need to balance the risks of miscarriage due to amniocentesis (approximately 1 in 200 procedures) with the benefits of identifying affected pregnancies.

It is important to stress that the majority of women identified as high-risk are in fact carrying a normal child, and not all Downs syndrome pregnancies are identified as high-risk. The evaluation of high-risk implies the woman has a higher likelihood than others to giving birth to a child with Down's and should be subject to further evaluation.

Management of Quality of Performance in Downs Screening

Like other NHS screening services, a Downs Screening Service has to maintain a quality system which works across the complete process. Checks are needed at every stage and this requires a positive multidisciplinary approach:

1. The initial discussion with the pregnant woman
2. The taking of the blood sample
3. Collection of robust demographic data
4. Assessment of foetal age
5. Maintenance of reliable transport arrangements for the sample
6. Provision of accurate laboratory analytical procedures
7. Use of reliable calculation algorithms and software

8. The delivery and clear presentation of the results by the relevant healthcare professional (midwife, nurse or doctor)

These checks are supported by several internal and external QA processes [Ferres et al., 2001].

5.1.3 Overview of error handling activities

During a period of five months, there were *three* different error recovery processes initiated by nurses in different locations; two from other hospitals, and one from another department within Sheffield Northern General. The first two attempts involved nurses that became increasingly concerned over a lack of high-risk reports coming back from the laboratory—they were however not certain that there was an error; this is considered as *problem detection* due to the uncertainty that comes with the concerns being raised. The first nurse noticed this discrepancy just two weeks after the bug came in effect, while the second made her first report two months later. From January to May, the two nurses made several reports to the Immunology Department, but they did not manage to convince laboratory staff that there was a problem with Down's screening. Reporting was primarily done over the phone, with different people picking up the phone on almost every occasion.

In May, five months after the bug manifested, the problem was realized almost accidentally by an investigation into—what seemed at the time—another error. More specifically, nurses from Antenatal care thought the dates of birth of two mothers had been wrongly entered in the system and requested that they be changed. When a

laboratory technician attempted to change the dates of birth, the risk calculation did not change; this raised some suspicion. Even when laboratory staff realized there was an error in Down's screening, they were not aware of the magnitude of the failure. In addition, communication breakdowns and absences of key personnel at that stage prolonged the incident even further (3-4 days). Eventually, when people decided to look into the software system, the bug was detected and fixed within 37 minutes.

Despite the resolution of the technical problem, a further investigation was necessary in order to identify and contact each affected pregnant woman. Until that investigation was finished, recovery could not be considered complete. Eventually, it was determined that 158 high-risk pregnancies had been labeled as low-risk before May 23rd. The women were contacted for reexamination, where two proceeded with a late abortion, and two other gave birth to a child with Down's syndrome.

5.1.4 ER-STEP analysis

This section will describe three different recovery processes, initiated by the

1. Maternity and Gynecology liaison sister, hospital B
2. Midwife coordinator, hospital C
3. Antenatal staff, Sheffield Northern General hospital

In addition, the activities that followed recovery and correction of the software errors regarding the evaluation of the impact of the failure will also be discussed.

Maternity and Gynecology liaison sister, hospital B

The first attempt to report concerns was made only two weeks after the bug began to manifest itself after the millennium. A Maternity and Gynecology liaison sister in another

hospital (Hospital B) was concerned there were not enough positive results, and made at least three attempts to indicate the error to the immunology department in January.

As the Downs Screening service had been offered for approximately 12 years, the nurse quickly became concerned by noticing a discrepancy in the frequency of high risk pregnancies diagnosed **[Event 1.4: problem detection]** . She reported her concerns to the lab over the phone in mid January **[Event 1.5: problem indication]**, but her reports were not seen as significant at the time (Figure 5.1 below illustrates the ER-STEP diagram that describes all the activities concerning the Maternity and Gynaecology liaison sister's efforts).

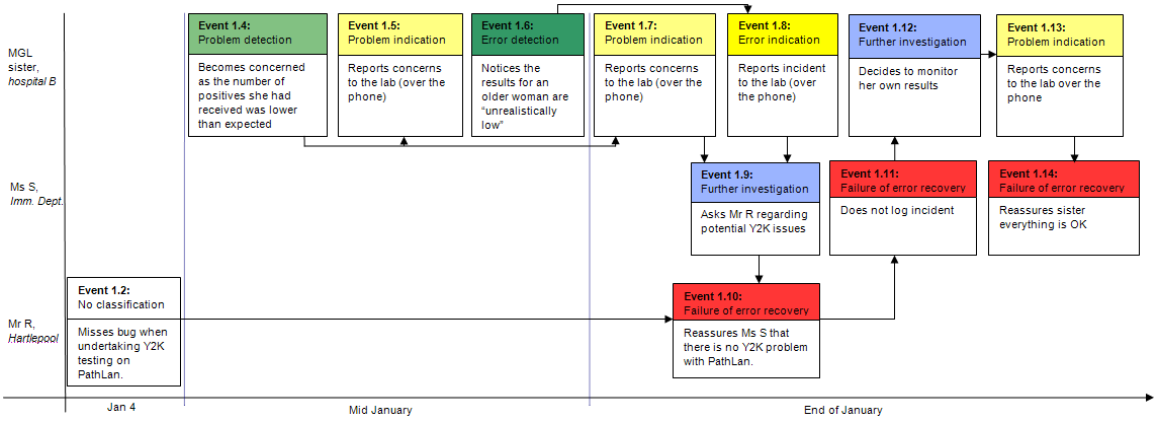


Figure 5.1: Error Recovery Efforts 1, Sheffield.

Shortly after she had reported her concerns regarding a lack of high-risk pregnancies, the same nurse was confronted with an incident that increased her concerns. An older woman was found to be in the low risk area, even though she should be in the high-risk area by default, because of her age[11] (all women over 35 are considered to be in high-risk) **[Event**

[11] This event is labeled as error detection rather than problem detection because an incident involving one specific diagnosis is a clearer indication of the error rather than general concerns of a lack of high risk pregnancies.

1.6: error detection]. The nurse reported the incident directly to the lab **[Event 1.8: error indication]**, and this time her concerns had as a result a brief inquiry **[Event 1.9: further investigation]**.

Ms S, the acting Medical Laboratory Scientific Officer (MLSO) at the time, informed a colleague of the incident and the nurse's concerns. 'Ms S' colleague' had performed the Y2K tests for the application **[Event 1.3: No classification]** that supported the Down's screening algorithm in the previous couple of weeks. During their phone conversation, Ms S's colleague reassured Ms S that everything was fine and that Y2K testing had been carried out correctly, and this ended the investigation **[Event 1.10: Failure of Further investigation].** As a consequence, Ms S did not consider the incident as serious, and decided not to log it in the 'high book', a book that is placed next to the phone for recording any reported abnormal results **[Event 1.11: Failure of Further investigation]** (It was suggested in the report that there was an overall high confidence in the Y2K compliance activities carried out in the lab).

Towards the end of January, the nurse was becoming frustrated as her reports were ineffective, and she decided to monitor results personally thereafter **[Event 1.12: further investigation].** She made her last attempt to notify the lab in the end of January **[Event 1.13: problem indication],** however she didn't make any progress. All reports were made over the phone. No more reference was made to the Maternity and Gynecology Sister since then in the inquiry report.

Midwife Coordinator, Hospital C

Approximately a month after the incident involving the Maternity and Gynaecology Liaison sister, the *Midwife Coordinator* of another hospital (Hospital C) also had growing concerns over a lack of high risk pregnancies **[Event 1.23: problem detection]**. From late April through to early May she reported her concerns to the lab twice over the phone **[Event 1.24: problem indication], [Event 1.26: problem indication]**. Her efforts had similar results with the Maternity and Gynecology Sister's, as there were no immediate actions to confirm or cancel out her concerns **[Event 1.24: failure of error recovery]**. In fact, at her second attempt, she spoke to Ms S **[Event 1.26: problem indication]** who had also been personally in contact with the Maternity and Gynaecology Liaison sister in January (**Events 1.7** and **1.8**); no association of the two separate reports was made by Ms S according to the inquiry report. In this occasion, Ms S suggested she would notify Mr M regarding this matter **[Event 1.27],** but there is no evidence in the report that she actually did **[Event 1.28]**, as she never responded to the Midwife's phone-call **[Event 1.29].**

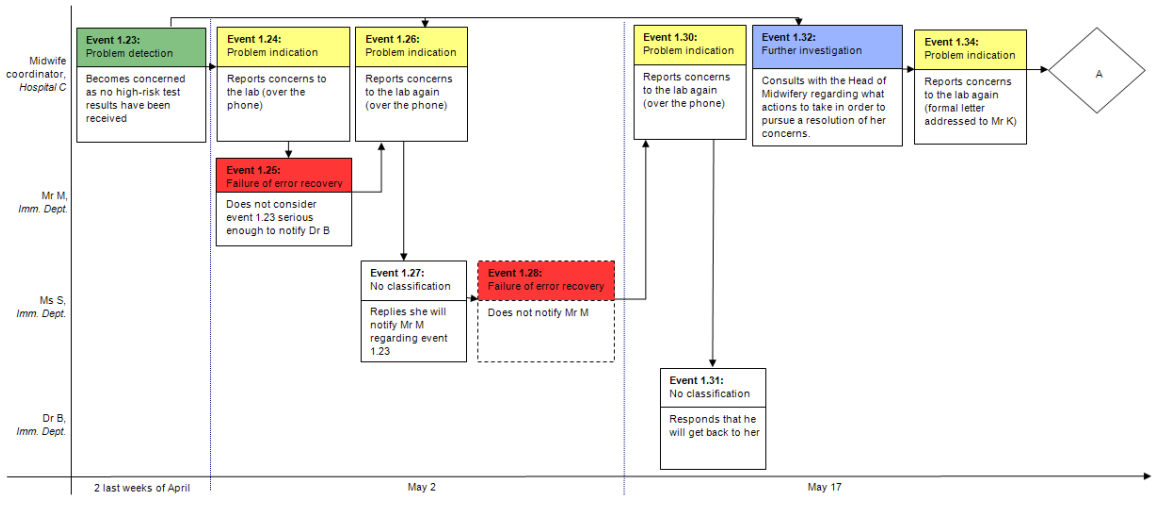


Figure 5.2: Error recovery efforts 2, Sheffield.

The Midwife Coordinator made another attempt to communicate her concerns to the lab over the phone **[Event 1.30: problem indication]**, and this time spoke to the Head of Department, Dr B. He told her he would get back to her soon **[Event 1.31: No classification].** In the meantime, the Midwife Coordinator consulted with the Head of Midwifery of Hospital C **[Event 1.32: further investigation]**, and decided to put her concerns in writing **[Event 1.34: problem indication]**. This decision was important and, as we shall see later on, it commenced a series of actions that led to the identification of the error; however, it followed a number of failed attempts to report the problem informally.

Later on during the same day (May 17th), Dr B discussed the matter with Ms P **[Event 1.35: further investigation]** another MLSO (apart from Ms S), who informed Dr B **[Event 1.36: further investigation]** that an audit which had been requested in April (Figure 5.3 illustrates the events that took place in April regarding the audit) and by the lab IT technician had not been carried out **[Event 1.37: No classification]**. Dr B instructed Mr M to have the audit report on his desk the following day **[Event 1.38: further investigation].**

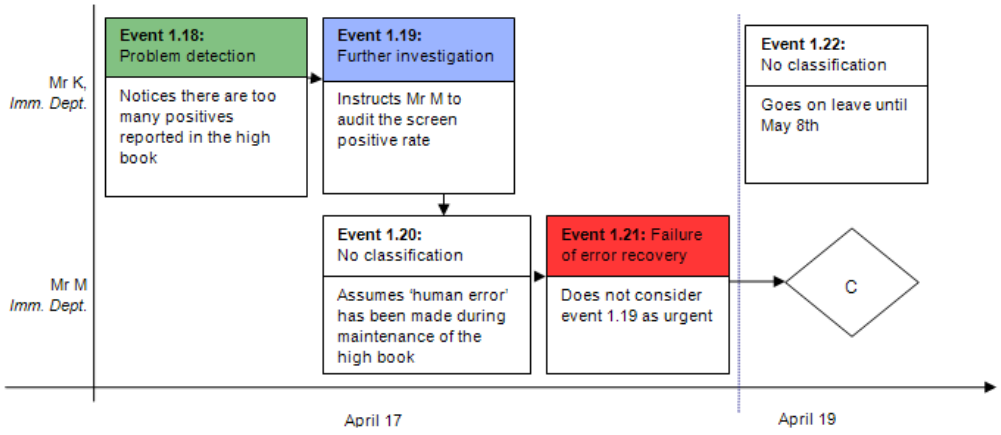


Figure 5.3: Error recovery efforts 3, Sheffield.

On May 18th Dr B saw the note on his desk **[Event 1.50: No classification]** but didn't realize the piece of paper was actually the result of the audit **[Event 1.51: Failure of Further investigation]**. It was suggested that at the time Dr B was very busy, while his wife was about to go into labor. The next day, Friday May 19th, the Midwife Coordinator made a phone call to the lab again reporting her concerns, and also to inform Mr K that she would be sending a letter to the lab about her concerns over the lack of high risk pregnancies **[Event 1.56: further investigation].** On that day Dr B's wife went into labor **[Event 1.53: No classification]**. During his absence, no further activities regarding Downs Screening took place until after the weekend.

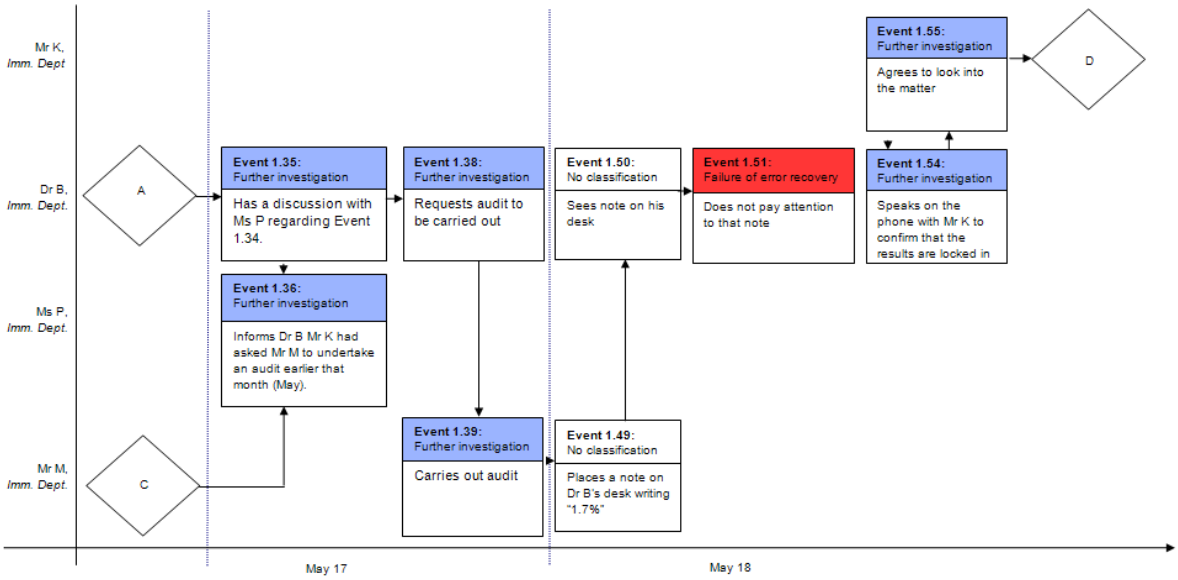


Figure 5.4: Error recovery efforts 4, Sheffield.

The Midwife Coordinator's reports did not have a direct impact in the detection of an error within the lab. However, during the third week of May, there was another report to the Immunology department which was seen at the time as a separate event. This report,

in conjunction with the efforts of the Midwife coordinator resulted in the realization of the problem with Down's screening.

Antenatal Staff, Sheffield Northern General

On May 17th, the same day that the Midwife Coordinator spoke to Dr B over the phone **[Event 1.42: error indication]**, staff from Antenatal care made a phone call requesting amendments to two reports. The dates of birth in the reports were wrong; one by a couple of months and one by a couple of years **[Event 1.41: error detection].** Note that error detection here refers to the detection of another error [error 2].

Ms S became concerned there was a problem with Downs Screening as when she changed the dates of birth accordingly, the risk calculation remained the same **[Event 1.44: error detection]**. Ms S reported this matter to Mr L, who was responsible for maintaining PathLan **[Event 1.45: error indication]**. Mr L then attempted to inform Mr W in Hartlepool **[Event 1.46: error indication]** who, along with Dr A, had created the Downs Screening software application back in 1988. Mr W was not available and so Mr L left a voice- message **[Event 1.48: error indication].**

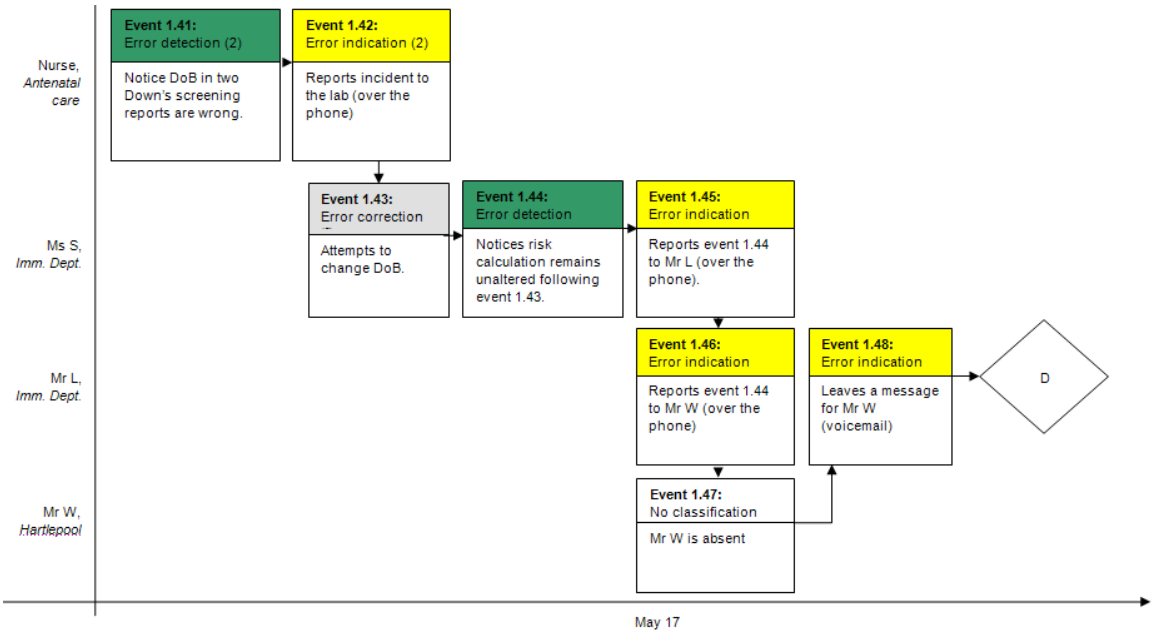


Figure 5.5: Error Recovery Efforts 5, Sheffield.

The following day, Friday 19th, Mr L was away, and no further actions were taken until Monday 22nd. On that Friday Mr K received the letter from the Midwife Coordinator **[Event 1.57: further investigation].** On the 22nd, Dr B's wife went into labour, so he was absent on paternity leave.

On May 23rd, Mr K asked Mr M to provide him with the audit results **[Event 1.60: further investigation]**. After examining the audit results **[Event 1.61: further investigation]**, Mr K found that high risk calculations were overall much lower than anticipated, and not just with regards to Hospital B's pregnant women, but to all recipients of their test reports [**Event 1.62: error detection]**. Mr K immediately asked his assistant Mr L to contact Mr W **[Event 1.64: error indication]** while he had already started checking analytical values himself **[Event 1.63: further investigation]**.

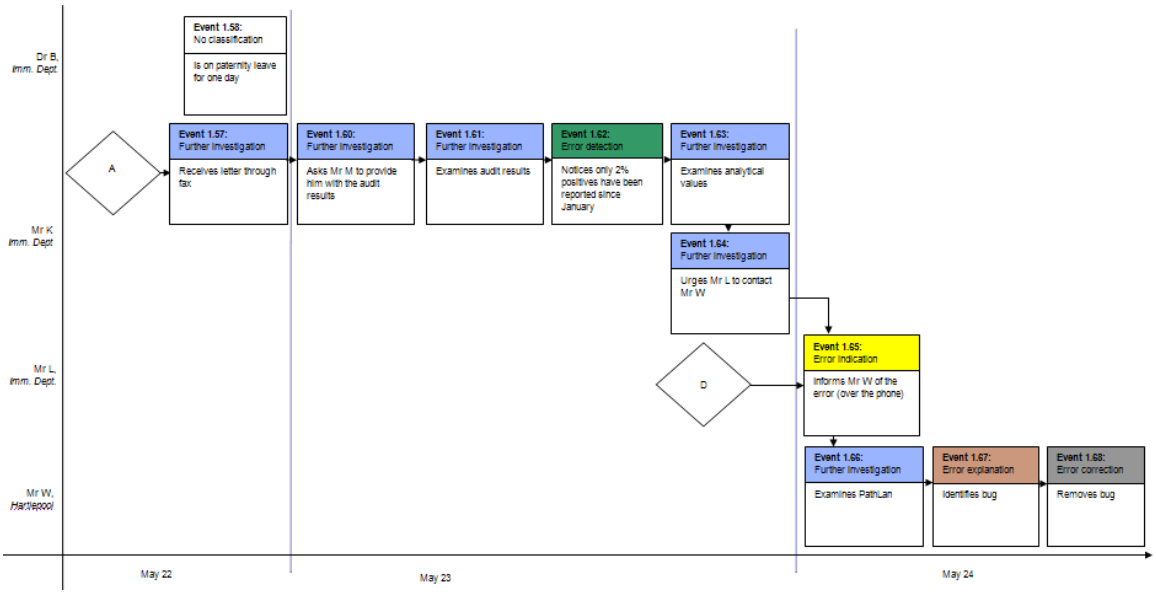


Figure 5.6: Error Recovery Efforts 6, Sheffield.

Contact with Mr W was established the next day **[Event 1.65: error indication]**. Mr W then logged on to PathLan in the morning and checked the date values **[Event 1.66: further investigation]**. Within 37 minutes, the bug had been identified **[Event 1.67: error explanation]** and corrected **[Event 1.68: error correction]**. Mr K did not find out about the Maternity and Gynecology Sister's reports until after the error had been corrected.

At that stage it was important to find out how many errors had been made. On the morning of May 24th, Mr K informed Dr B that he would find all high-risk cases that had been wrongly reported. He trawled through the system to identify the potential size of the problem over the next 12 hours or so **[Event 1.70: further investigation]**. He found approximately 150 high-risk pregnancies which had been reported as low-risk **[Event 1.71: error explanation]**—Mr K subsequently emailed these findings to Dr B **[Event 1.72: further investigation]** who then informed the Chief Executive.

5.1.5 Further analysis

As mentioned in Section 4.3, the final stage of the ER-STEP analysis is to rearrange the sequence of events based on the progression that is suggested by the error recovery framework. With this view, the analyst can focus even further on the process of error recovery.

In this section, the ER-STEP diagrams will therefore be rearranged to the error recovery view, while detection events will be classified according to the different detection types proposed by Sellen [1994] (action-based, outcome-based, through limiting functions). These findings will be aggregated and further analyzed in the next chapter.

Maternity and Gynecology Liaison Sister, hospital B

Figure 5.7 illustrates the same events as in Figure 5.1 with the error recovery view.

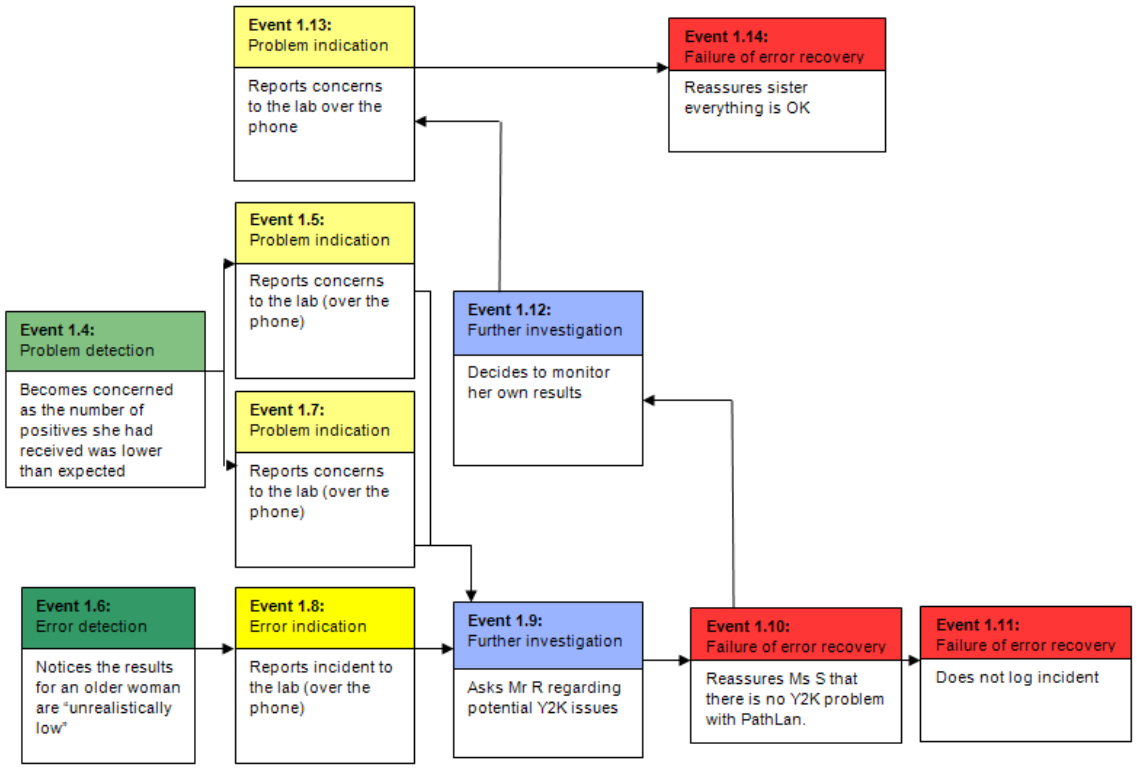


Figure 5.7: Error recovery focused view, Sheffield 1.

As we can see, detection events occurred where test results are used and not inside the lab. Reporting of either concerns or of an incident with a patient had to be reported to the lab over the phone several times. In the first instance, problem indication did not convince laboratory staff that there was actually a problem. Following several reports by the same person, there were some investigation activities—although these came close to the identification of the software bug, the person who had carried out Y2K testing dismissed these reports, so the incident was not investigated further and was not logged.

Table 5.2 summarizes the different kinds of activities under each stage that have taken place during the efforts initiated by the Maternity and Gynaecology sister during January 2000.

Event	Analysis
Problem detection	**Outcome-based, outside the lab** Growing concern due to a discrepancy between the expected number of positives and the actually received (event 1.4).
Error detection	**Outcome-based, outside the lab** Outcome (screening result) different than what she expected (event 1.6).
Problem indication	Over the phone (events 1.3, 1.5 and 1.7)
Error indication	Over the phone (event 1.8)
Further investigation	**Outside the lab** Monitoring of results to find out if there is an error (event 1.12) **Inside the lab** Enquiry to colleague to find out if there is an error (event 1.9) Investigation to determine the extent of the failure (event 1.70)
Failure of error recovery	Incident not logged (Event 1.11) Reassurances that 'everything is OK' (events 1.10, 1.14)

Table 5.2: Analysis of recovery events, Sheffield 1.

Midwife Coordinator, Hospital C

Figure 5.8 illustrates the error recovery focused view of the activities following the Midwife coordinator's efforts presented in Figure 5.2 and Figure 5.4.

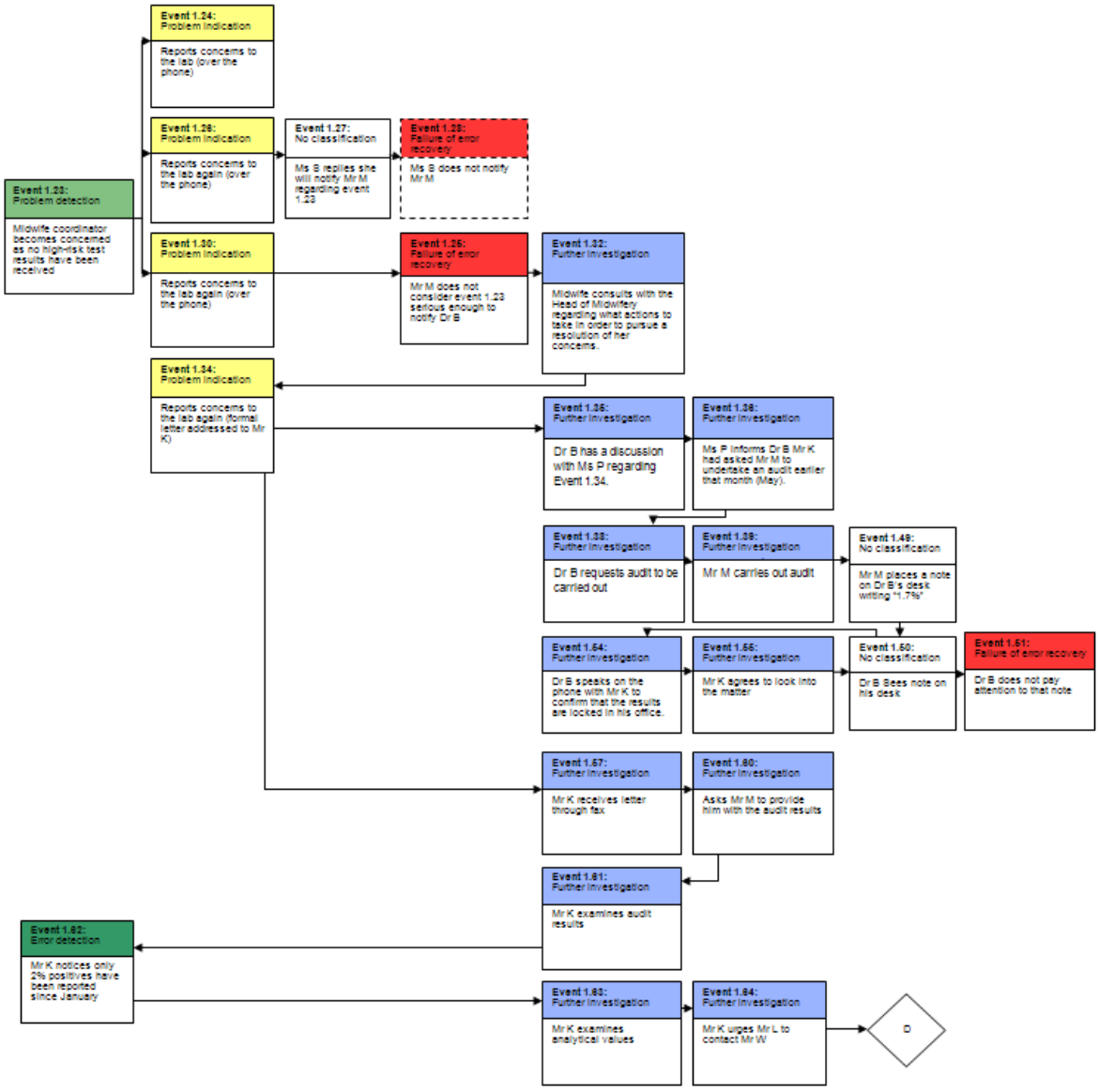


Figure 5.8: Error recovery focused view, Sheffield 2, 4 and 6.

It is obvious from the above figure that the reports made by the Midwife coordinator resulted in a significant amount of investigation activities. There were at least five distinct events where staff was investigating the system in order to determine if there is an error, after people in the laboratory started becoming suspicious. However, this was largely motivated by the Midwife coordinator's decision to write a formal letter, while her phone-call to the laboratory to announce she had in fact sent that letter perhaps accelerated investigation activities in the laboratory.

However, these events were not enough to lead laboratory staff to identify the actual cause of the errors (although it makes sense to assume that investigation activities that were undertaken at the time would have resulted in *error explanation* anyway). This was done when staff from antenatal found errors which seemed at the time as unrelated—this will be discussed in the following section.

Table 5.3 summarizes the different kinds of activities under each stage that have taken place during the efforts initiated by the Midwife coordinator in Hospital C during March-May 2000.

Event	Analysis
Problem detection	**Outcome based, outside the lab** Growing concern due to a lack of positives reported (event 1.23)
Error detection	***Through further investigation*, inside the lab** Investigation following problem detection has resulted in the identification of an error (event 1.62)
Problem indication	Over the phone (events 1.24, 1.26 and 1.30) Written (event 1.34)
Error indication	None
Further investigation	**Outside the lab** Enquiry to colleague regarding what action to take (events 1.32) **Inside the lab** Enquiry to colleague regarding what action to take (events 1.35) Enquiry to colleague regarding actions that have been performed in the past (events 1.36,1.54 and 1.57) Investigation to find out if there is an error (events 1.38, 1.39, 1.55, 1.61 and 1.63)
Failure of error recovery	Failure to consider previous event as important (events 1.25, 1.28 and 1.51)

Table 5.3: Analysis of recovery events, Sheffield 2, 4 and 6.

Note that *error detection through investigation* is a detection mechanism which is not considered in Sellen's taxonomy. This kind of detection occurred following problem detection, problem indication and consequently, further investigation.

Antenatal Staff, Northern General

Figure 5.9 presents the error recovery focused view of events that are described in Figure 5.5 and Figure 5.6.

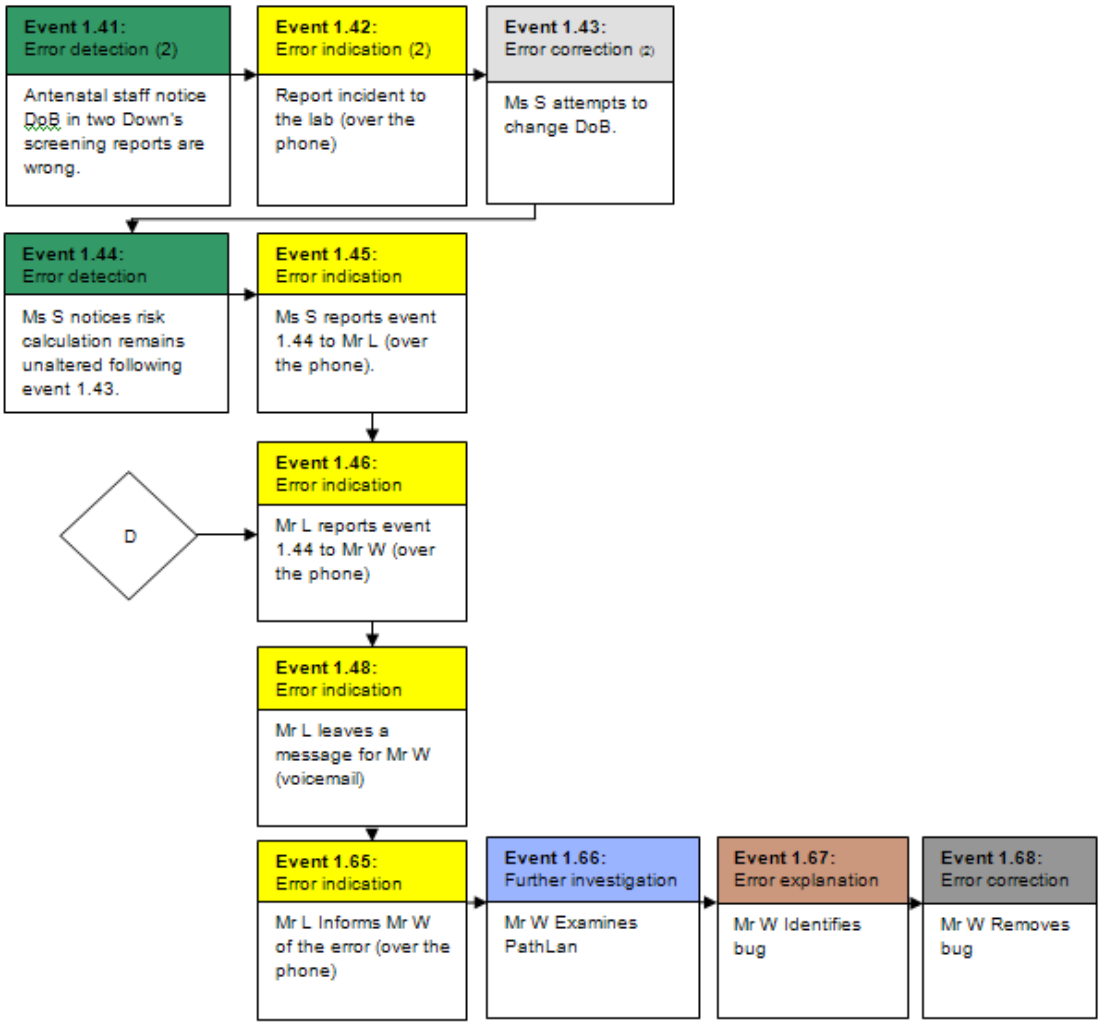


Figure 5.9: Error recovery focused view, Sheffield 5 and 6.

During a rather turbulent time for the Department (recent change of Head of Department, low staffing levels), with the reports from the Midwife coordinator under investigation, the two nurses in antenatal care found two reports with the DOBs of two women being wrong. When they phoned the laboratory to have the dates changed, it was realized that the risk calculation remained the same, which made it obvious that there was something wrong with the Down's screening software. It is important to stress that the person who tried to change the date was Ms S, who had been involved in all previous reports; she was aware of the reports by the Gynaecology and Maternity Liaison sister and the Midwife coordinator.

Mr W, the person responsible for the maintenance of the software was finally contacted that time—according to the report, it took him 37 minutes to identify and correct the bug.

Table 5.4 summarizes the different kinds of activities under each recovery stage that have taken place during the efforts initiated by Antenatal care staff towards the end of the incident in May 2000.

Event	Analysis
Problem detection	None
Error detection	**Outcome based, outside the lab** Antenatal staff found two reports where the dates of birth (DoBs) were wrong. This was an error during data entry, and could have happened regardless of the software bug. For this reason it has been labeled as error 2. (Event 1.41). **Outcome based, inside the lab** In an attempt to correct error 2, Ms S found the risk calculation remained the same. This led to the realization that there was something wrong with the risk calculation software (event 1.44)
Problem indication	None
Error indication	Over the phone (events 1.45, 1.46, 1.48) and voice-mail (event 1.65)
Further investigation	Investigation of system to find out what the errors is (event 1.66)
Error explanation	Identification of software bug (event 1.67)
Error correction	Removal of bug (event 1.68)

Table 5.4: Analysis of error recovery activities, Sheffield 5.5 and 5.6.

5.1.6 Overview of findings

Table 5.5 summarizes and categorizes all types of error recovery activities that took place during this incident.

Event	Analysis
Problem detection	**Outcome based, outside the lab** Growing concern due to a lack of positives (events 1.4, 1.23)
Error detection	**Outcome based, outside the lab** Older woman should be in high-risk by default (Event 1.6). **Outcome based, inside the lab** When trying to change a DOB, risk calculation remains the same (event 1.41) **Through investigation, inside the lab** Previous reports investigated result into identification of error (event 1.62)
Problem indication	Over the phone (events 1.3, 1.5, 1.7, 1.24, 1.26, 134)

Event	Analysis
	Written (event 1.34)
Error indication	Over the phone (events 1.45, 1.46, 1.48) and voice-mail (event 1.65)
Further investigation	**Outside the lab** Monitoring of results to find out if there is an error (event 1.12) Enquiry to colleague regarding what action to take (events 1.32) **Inside the lab** Enquiry to colleague to find out if there is an error (event 1.9) Enquiry to colleague regarding what action to take (events 1.35) Enquiry to colleague regarding actions that have been performed in the past (events 1.36,1.54 and 1.57) Audit to find out if there is an error (events 1.38, 1.39, 1.55, 1.61 and 1.63) Investigation of system to find out what the errors is (event 1.66) Investigation to determine the extent of the failure (event 1.70)
Error explanation	Identification of software bug
Error correction	Removal of bug
Failure of error recovery	Incident not logged (Event 1.11) Reassurances that 'everything is OK' (events 1.10, 1.14) Failure to consider previous event as important (events 1.25, 1.28 and 1.51)

Table 5.5: Summary of error recovery activities in the Sheffield incident.

5.2 Case study 2: Breast cancer screening errors, London

This second incident is particularly different from the first one. The error was not originating in software or in the analysis altogether, but in the notation laboratory technicians used to denote positives. The notation was "confusing", leading to multiple positives missed. In addition, there was no protocol in place to ensure that women were receiving the correct results. In comparison to the first incident, there are more organizational deficiencies directly involved with the manifestation of diagnostic errors.

In addition, this incident was severely prolonged by the subsequent investigations that were carried our.

However, the impact of the temporal dimension on the clinical outcome is not as immediate as in the first case study which considered diagnoses on pregnancies. The impact in terms of numbers of patients was therefore smaller, although fatal in some cases.

5.2.1 Incident summary

In mid October 2000, it was discovered almost accidentally that a woman had been sent the wrong results (she was informed she was fine, although she should have been called for further testing) by her previous breast screening service following her mammogram in January 1999. Following this incident, two inquiries were carried out. During this investigation, over 174,000 screening episodes were reviewed, concluding that 123 women had not received the right result. The error was eventually associated with a delay in diagnosis of breast cancer in 11 women, while the longest delay was 21 months. One woman's condition deteriorated and she died.

Incident	Errors in Breast Cancer Screening service
Incident timeframe	1993 – December 2001
Primary cause	Absence of protocol to ensure women receive the correct results
Data source	*Formal inquiry report* [CHI, 2002]

Table 5.6: Overview of incident 2.

One of the conclusions drawn by the inquiry committee was that: "…*there had been warning signs of the potential for service failure which had not been reported and, therefore, not acted on by West of London Breast Screening Service…"*

5.2.2 Background

The inquiry report for this incident presents only the events that followed the detection of the error, and mostly the activities surrounding the decisions and conduct of three inquiries that followed (one internal, one independent, and a formal NHS inquiry—the latter being the source of this analysis). Errors were made over a period of eight years; they were not attributed to software or hardware faults, but to a lack of protocol to ensure that the right results were given to the right patient.

There were two Breast Cancer Screening services involved: the West of London Breast Screening Service (WLBSS), where the errors occurred, and Breast Screening Service X (BSS X), where one of the errors was detected. Abbreviations here are the same as in the formal inquiry report.

5.2.3 ER-STEP analysis

In January 1999, a woman was sent the wrong mammogram tests by the WLBSS **[Event 2.1: No classification]**, where she was cleared from any risks associated with Breast Cancer. She should have been called back for further testing. In October 2000, she moved to another area **[Event 2.2: No classification]** and had her files sent over to her new Breast Cancer Screening Service, BSS X **[Event 2.3: No classification]**. A review of her

case found the error **[Event 2.4: error detection]**, and the BSS X contacted the WLBSS over the phone regarding the matter **[Event 2.5: error indication]**.

On October 31st, the BSS X sent written confirmation to the WLBSS **[Event 2.6: error indication]**, that there was an error with the specific patient's diagnosis, and copied the letter to their own Quality Assurance Centre (Quality Assurance Centre X). The matter was discussed among senior management within WLBSS **[Event 2.7: further investigation]**, however it was not regarded as significant **[Event 2.8: Failure of error recovery]**, as there had been no complications for the woman's health.

On the following day, BSS X informed the London Quality Assurance Reference Centre **[Event 2.9: error indication]** over the phone, while they also forwarded the letter sent to the WLBSS the previous day **[Event 2.10: error indication]**. The letter was received 10 days later (November 10) by the London Quality Assurance Reference Centre **[Event 2.11: further investigation]**. Until then, no further actions to investigate the error and other possible complications were carried out.

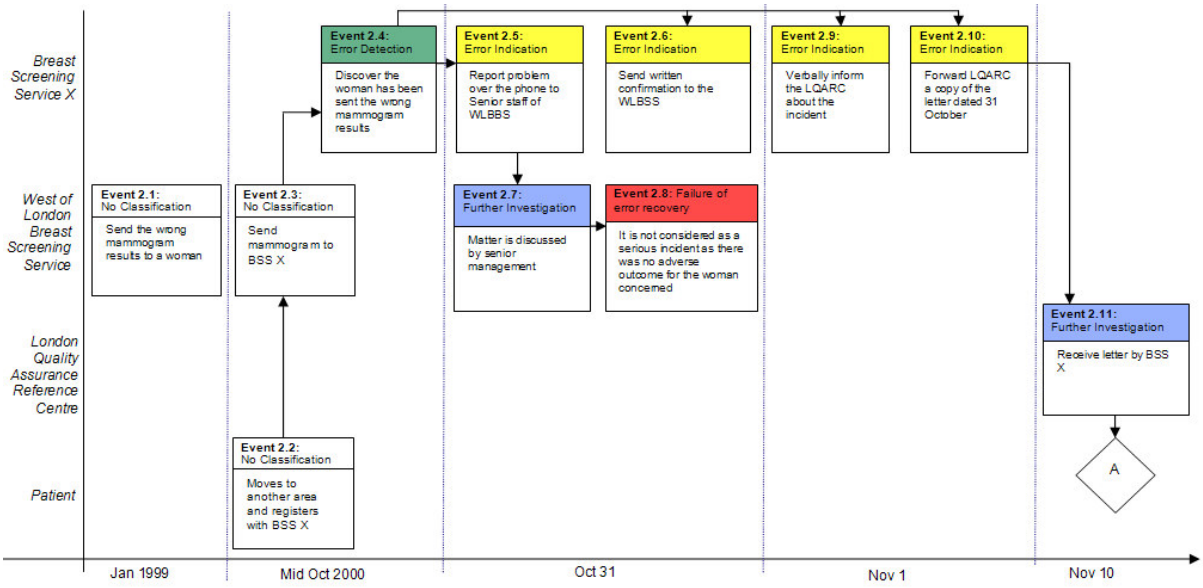


Figure 5.10: Error recovery efforts 1, London.

On the week starting November 13th, the London Quality Assurance Centre made several attempts to discuss the matter with senior management of the WLBSS **[Event 2.12: error indication]**. They insisted the error be reported to the General Manager of the Hammersmith Hospitals Trust. On November 17th, the incident was finally reported as a 'critical incident' to the Trust's General Manager **[Event 2.13: error indication]**, who immediately contacted the Trust's Chief Operating Officer **[Event 2.14: error indication]**. 5 days later, the matter was brought to the attention of the NHS England Coordinator of Breast Screening **[Event 2.15: error indication]**.

The following day (Nov 23), the London Quality Assurance Reference Center informed the officer with lead responsibility for Cancer Services at the NHS London Regional Office **[Event 2.16: error indication]**. On November 24, the London Quality Assurance Reference Centre wrote a letter to the NHS Region Director of Public Health regarding the incident **[Event 2.17: error indication],** and decided to call a meeting **[Event 2.18:**

further investigation], with participants being representatives from the WLBSS, the wider trust, the health authority commissioning consortium, and the Quality Assurance Reference Center, while the trust's Chief Operating Officer was chairing the meeting which was scheduled for December 4th.

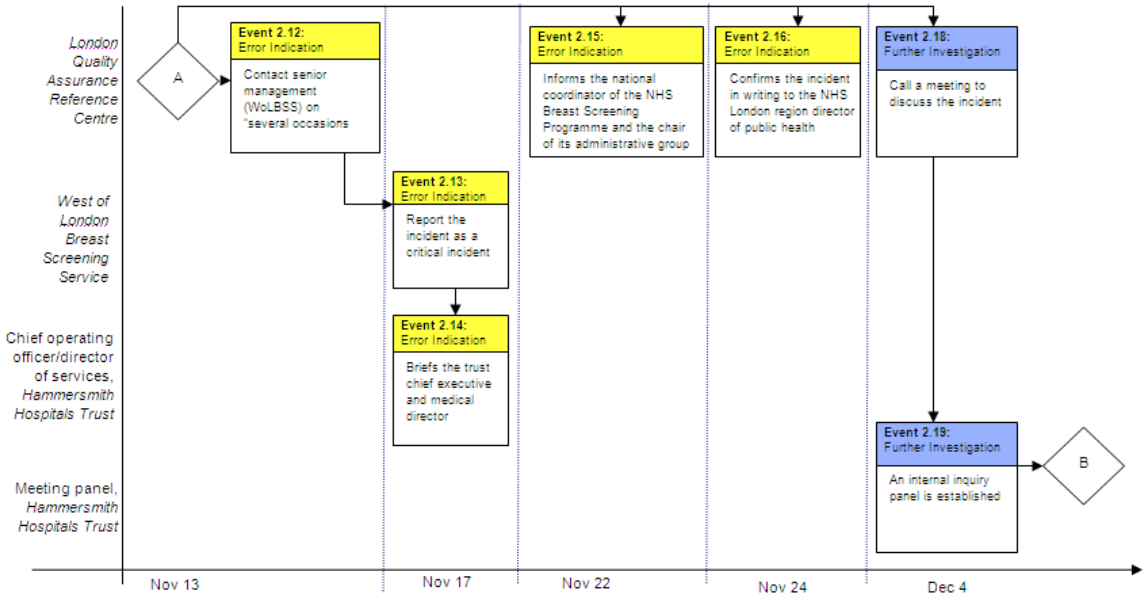


Figure 5.11: Error recovery efforts 2, London.

At the meeting, the trust established an *internal inquiry panel* **[Event 2.19: further investigation]**. The panel, made up by the general manager of the directorate and a consultant radiologist who did not work at the WLBSS, reviewed a number of documents and conducted interviews **[Event 2.20: further investigation]**. It was understood that the WLBSS needed to develop a robust right results protocol to ensure women received the correct result **[Event 2.21: error explanation]**, while they suggested an external audit company should review the mammogram files of all women who had attended for screening since 1993, nearly 104,000 women (over 174,000 episodes) **[Event 2.22: further investigation]**.

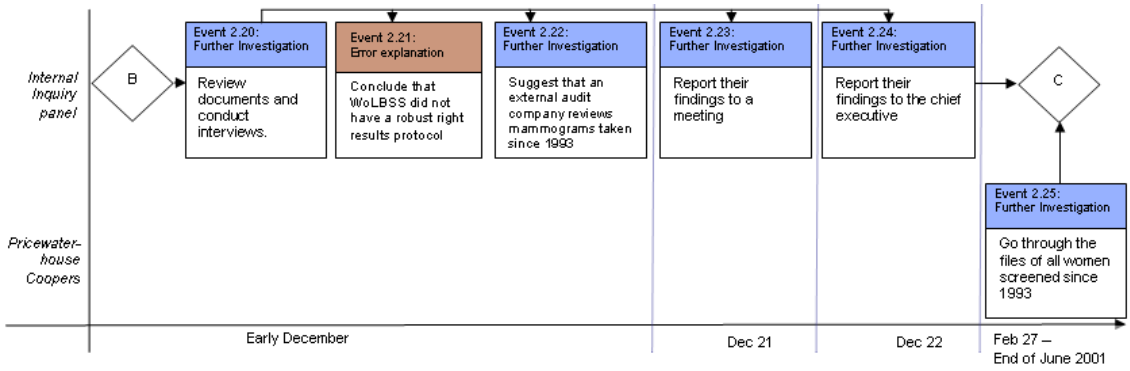


Figure 5.12: Error recovery efforts 3, London.

The findings of the inquiry panel were reported to the Trust Chief Executive on December 22nd **[Event 2.24: further investigation]**. The Trust Board also discussed the matter in a closed session, on February 12th 2001 **[Event 2.23: further investigation]**. It was decided an external audit should be carried out, as suggested by the inquiry panel. The company—PricewaterhouseCoopers—was instructed to go over all the files since 1993, in order to identify any difference between the information contained in WLBSS files and the corresponding computer records. The principal objective was to identify cases in which women screened by WLBSS may have received the wrong result and incorrectly refereed for a routine recall in 3 years time instead of being recalled for immediate clinical or technical assessment.

This audit lasted approximately 3 months **[Event 2.25: further investigation]**; the findings were forwarded to the inquiry panel, which compiled a report in July **[Event 2.26: Error explanation]**. During the audit by PricewaterhouseCoopers, the Trust Chief Executive requested the assistance of the Commission for Health Improvement (CHI) [**Event 2.27: error indication]**, who agreed to conduct their own investigation on April 10th **[Event 2.28: further investigation]**.

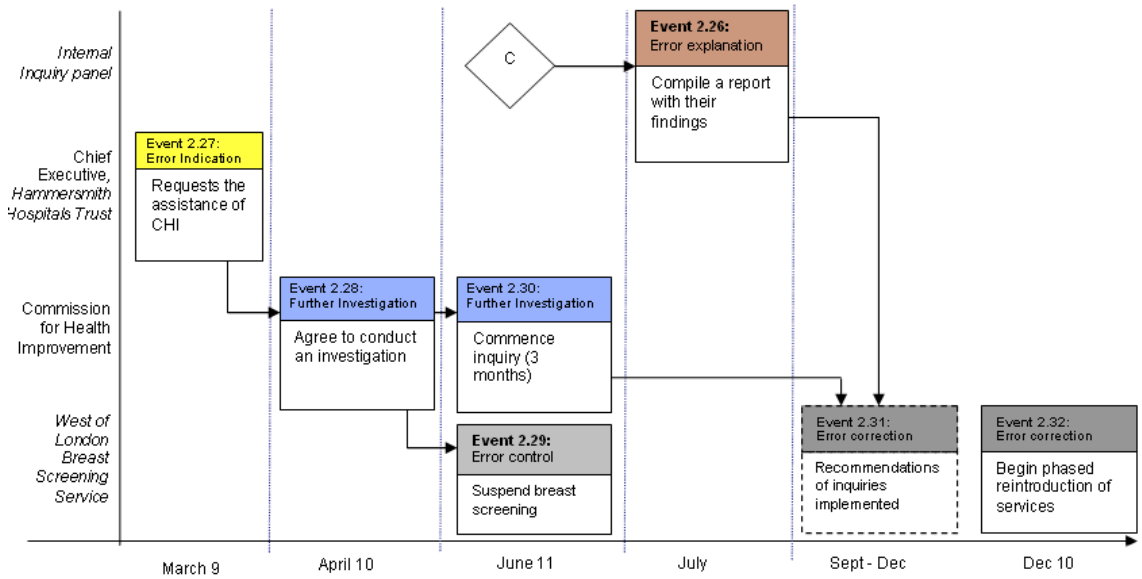


Figure 5.13: Error recovery activities 4, London.

On June 11th, WoLBSS suspended its breast cancer screening service **[Event 2.29: error control],** and CHI began their investigation **[Event 2.30: further investigation]** which was completed in April 2002. It was not until December 10th that WoLSBSS began a phased reintroduction of services **[Event 2.31: error correction]**.

5.2.4 Further analysis

Figure 5.14 presents the error recovery focused view of the error handling activities that took place during the Breast Cancer screening errors, London.

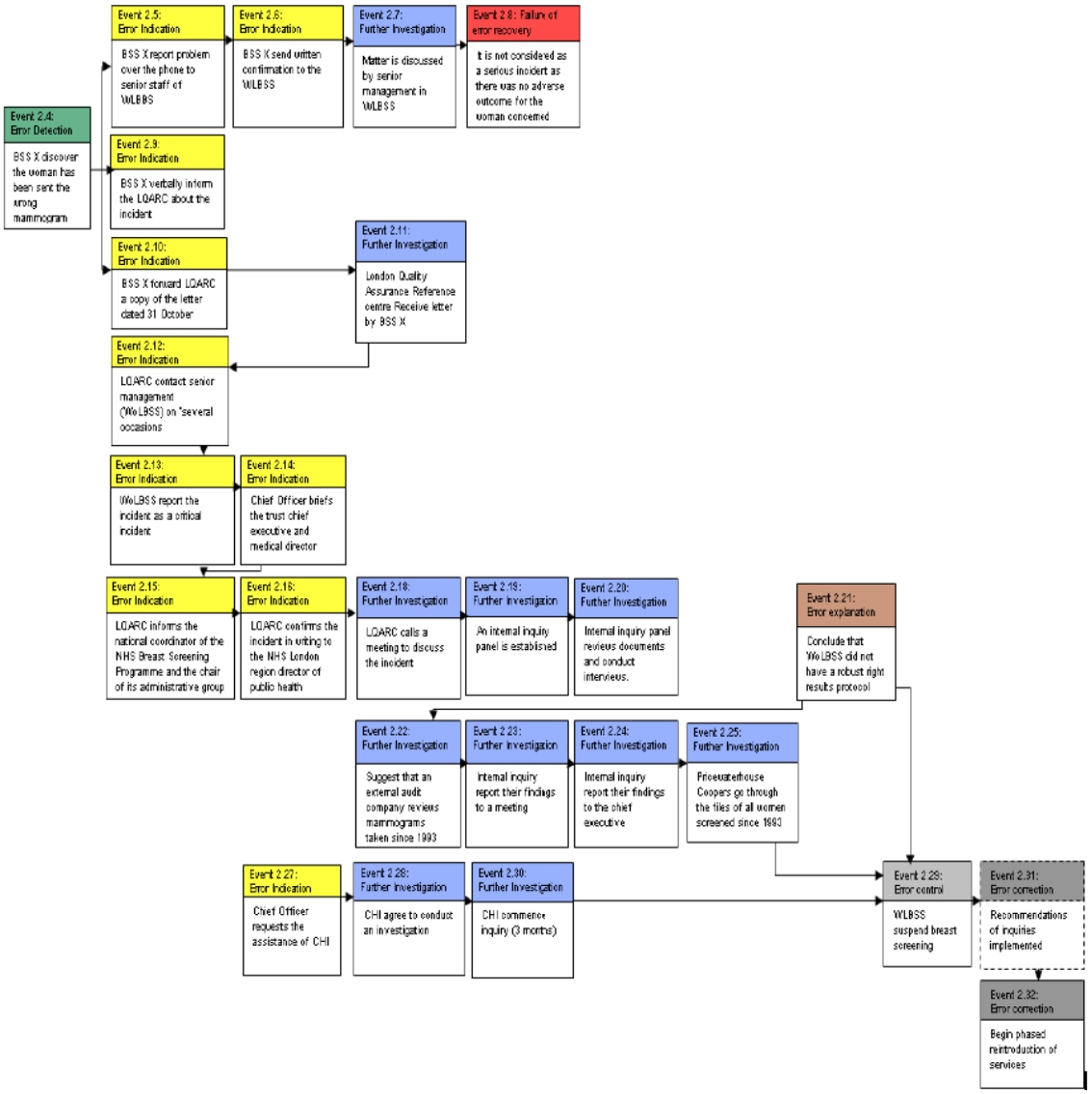


Figure 5.14: Error recovery focused view, London.

Error recovery efforts for this incident were triggered by a single instance of error detection. This was however done by a Breast Screening Service and not by a nurse in points of care. The reason for the extended recovery timeframe is the fact that errors were taking place for over eight years. A total of three inquiries had to be carried out in order to determine the number of erroneous reports. These inquiries are considered as part of the entire recovery process as the errors were continuing to affect patients as time was passing.

There also many error indication events, where the Breast Screening service that detected the error had to notify several agencies and authorities in order to establish a formal inquiry.

Table 5.7 summarizes and categorizes all error recovery events that took place during this incident.

Event	Analysis
Problem detection	None
Error detection	**Outcome based, outside the lab** Revaluation of a patient's screening result by new breast screening service (Event 2.4)
Problem indication	None
Error indication	Over the phone (events 2.5, 2.9 and 2.12) Formal letters (2.6 and 2.10) Incident reporting to NHS authorities (events 2.14, 2.15 and 2.16)
Further investigation	**Outside the lab** None **Inside the lab** Meeting to decide what action to take (event 2.18) Investigation to determine what the error is (events 2.19, 2.20, 2.28 and 2.30) Audit to determine the extent of the error (event 2.22 and 2.25)
Error explanation	Identification of cause (event 2.21)
Error control	Suspend breast screening (event 2.29)
Error correction	Implement recommendations (event 2.31), phased introduction of services (event 2.32)
Failure of error recovery	Previous event not considered as important (event 2.8)

Table 5.7: Summary of error recovery activities in the London incident.

5.2.5 A case of whistle-blowing?

Only a few days after the CHI report was published in April 2002, two former employees of the Hammersmith Hospitals NHS Trust gave an interview to BBC news, stating that there were further failures in the Radiology Department of the Trust[12]. In the interview, the one former employee argued that she—as well as other colleagues—had made several attempts to report problems in the radiology unit, which had to do with technology and management failures, but were not affective.

One of them suggested that "*Many scans were rendered unusable because reports were generated with missing characters and lines and even patients' names transposed*". The same person went on to argue that: "*Many compromised reports were simply abandoned because it was impossible to identify who the patient was, while even reports which were identifiable were abandoned because staff were under time pressure*".

These events took place during 1993-1995, when the nurse having written a letter to the management reporting these issues was suspended from her duties (allegedly within two hours). In 1995, A Picture Archiving and Communication System (PACS)[13] was introduced, which significantly changed the way with which things were done at the Radiology unit. However, other mishaps had taken place since 1999 according to the

[12] More information can be found at:
http://news.bbc.co.uk/1/hi/programmes/newsnight/archive/1938095.stm

[13] PACS systems were discussed in section 2.2.1

second whistle-blower. In fact, that person suggested that staff involved believed the inquiry carried out was incomplete, and even acted as a 'cover-up' of the problems that existed in the radiology unit and the breast cancer screening service.

The press article that discusses these matters is neither an official document, nor does it contain enough information for an analysis. However, it is worth mentioning as such occurrences illustrate the potential friction that may arise when staff in lower levels of the organization have concerns about the practices of their department or unit; the following quote from the involved nurse may highlight the challenge of whistle-blowing: "*I could never say to another person who might be in the NHS now, possibly watching this, thinking perhaps I should blow the whistle, I couldn't tell them go ahead and do it*".

5.3 Case study 3: Breast cancer screening errors, Manchester

This incident was attributed to 'human error' of a single radiologist. In comparison to the two previous case studies, the direct cause was therefore also different. However, like the previous cases, error recovery was poor and contributed to having a prolonged incident timeframe.

5.3.1 Incident summary

Over a two-year period, a consultant radiologist misinterpreted a total of 176 mammograms. 28 of these had previously been cleared by the radiologist, but were eventually identified as having breast cancer, out of which 17 were given reduced chances of survival. Although the radiologist involved was initially considered solely responsible, the investigations that followed concluded that severe organizational and

structural problems allowed for the errors to occur. The radiologist was, at the time working at two NHS Trusts: The Trafford Hospitals Trust and The Bury Primary Care Trust.

Incident	Errors in Breast Cancer Screening service
Incident timeframe	April 2003 – January 2006
Primary cause	'Human error'
Data source	*Two Formal inquiry reports* [Baker, 2006] and [Expert Advisory Panel, 2006]

Table 5.8: Overview of incident 3.

Two inquiries were carried out ([Expert Advisory Panel, 2006] and [Baker, 2006]). The two inquiries discussed here had different purposes: The Baker report considered the practices of the radiologist held responsible (Dr H), while the Expert Advisory Panel's report focused on the communications, meetings and reviews that were carried out upon the discovery of the errors.

The inquiry reports suggested that the errors would have been identified sooner had audit arrangements been in place as recommended in previous reviews, while the problems "*…may well have been masked in previous settings by the strength of their imaging department and of the breast multi-disciplinary team"*. In addition, it was concluded that "*…warning signs were missed or ignored and inadequate attention was paid to the nature of references"*.

5.3.2 ER-STEP analysis

Concerns about Dr H's practice were raised from the first couple of weeks in his appointment. More specifically, other Mammography radiologists were concerned that Dr

H's reports were too short, while he didn't take into account previous radiology reports **[Event 3.1: problem detection]**. Concerns were higher with respect to cancer patients under surveillance. However, no errors had been identified until then. Radiologists formally reported their concerns about Dr H's practices in November 2003 **[Event 3.2: problem indication].** The subsequent investigation **[Event 3.3: further investigation]** did not find any significant problems with Dr H's work **[Event 3.4: failure of error recovery].**

During November 2003, some errors in Dr H's reports were noticed by other radiologists **[Event 3.5: error detection]**—we have no information regarding the nature of the errors. Trust management became aware of this **[Event 3.6: error indication]**, but the errors were seen as isolated events, and not systematic **[Event 3.7: failure of error recovery].**

During November, clinical staff at Trafford were also becoming concerned that Dr H's work was not reliable **[Event 3.8: problem detection]**. As a result, they were checking all critical tests with another radiologist **[Event 3.9: further investigation]**. However, they were not checking all results, but only the ones with diagnosed cancer (therefore only consider the risk of false positives). Since then, the inquiry reports do not mention any further activities with regards to Dr H's work until April 2005.

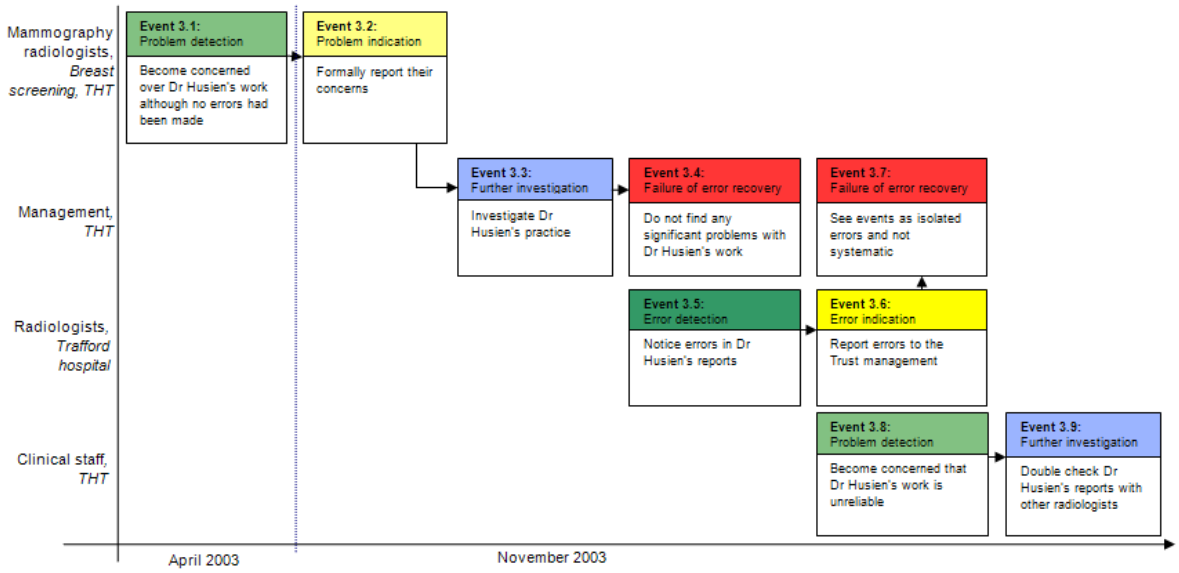


Figure 5.15: Error recovery efforts 1, Greater Manchester.

In April 2005, mammography radiographers at Trafford Hospitals Trust found a higher number of errors than expected in a single Breast Care Multi-Disciplinary Team (MDT) patient list **[Event 3.10: error detection].** This, along with their general concerns regarding Dr H's work was reported to the THT management **[Event 3.11: problem indication], [Event 3.12: error indication].** On April 18th, the Trust management decided to suspend Dr H **[Event 3.13: error control].** The following day, THT management reported Dr H's errors as a 'serious adverse event' to the Greater Manchester Strategic Health Authority (GMSHA) **[Event 3.14: error indication]**, who subsequently informed the Department of Health **[Event 3.15: error indication].** An independent review was called, and the Nightingale centre was instructed to conduct it **[Event 3.16: further investigation]**.

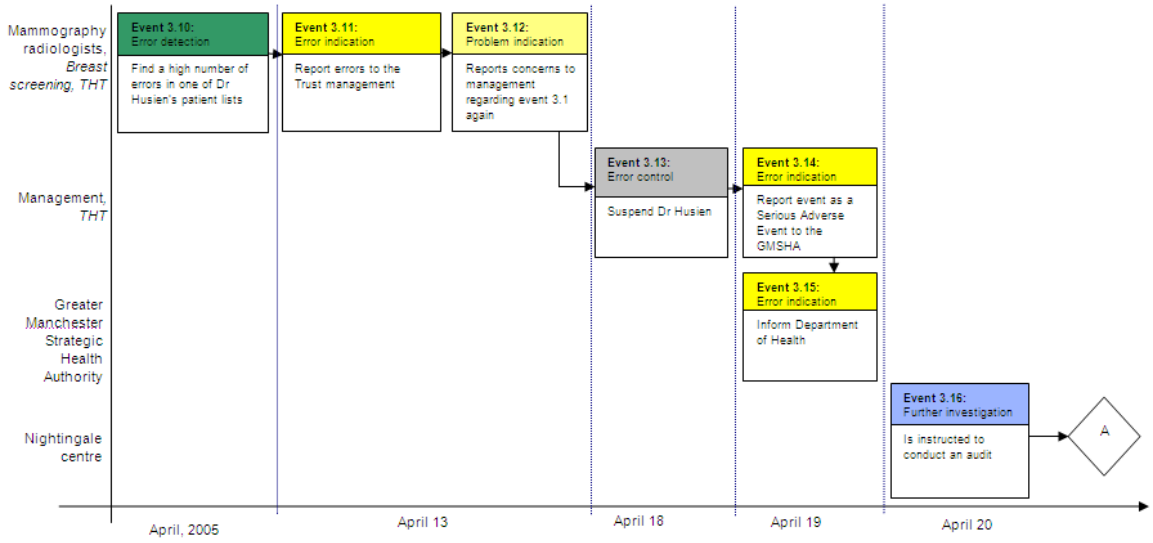


Figure 5.16: Error recovery efforts 2, Greater Manchester.

On April 26th, the Nightingale centre commenced a mammography review (consisting of 478 reports) **[Event 3.17: further investigation]**. On the same day, an Expert Advisory Panel was established, in order to advise the Trusts on the management of the clinical incident **[Event 3.18: further investigation].** The mammography review by the Nightingale centre was concluded on May 6th, finding a significant number of differing reports **[Event 3.19: error explanation]**. At that stage, the Bury Primary Care Trust (where Dr H was also working on a part-time basis) was advised to exclude Dr H **[Event 3.20: error indication]**. The National Patient Safety Agency was informed of the incident on April 17th **[Event 3.21: error indication]**.

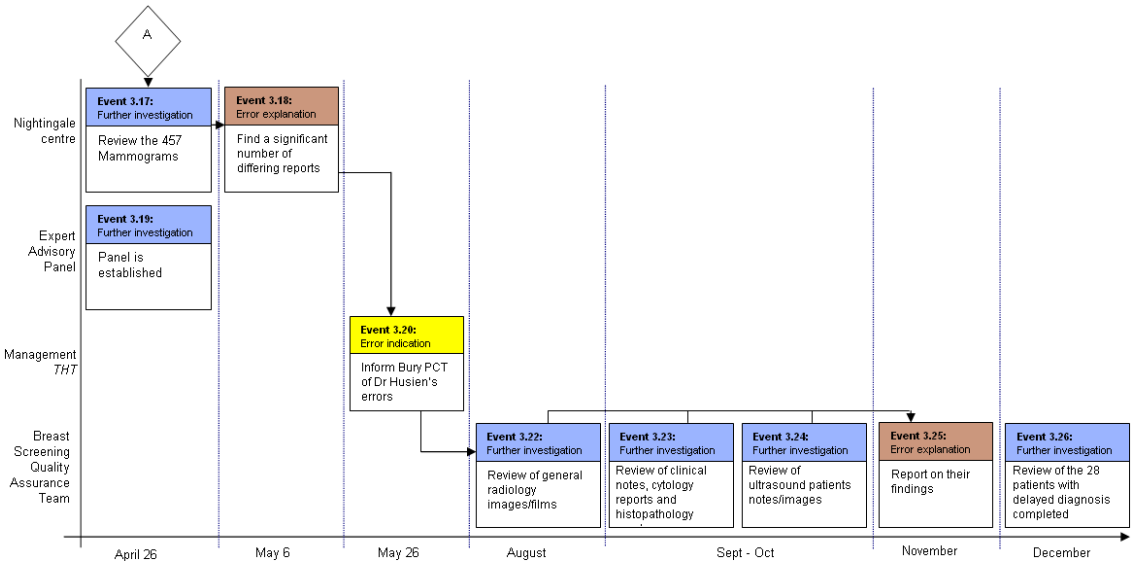


Figure 5.17: Error recovery efforts 3, Greater Manchester.

During August, the London Breast Screening Quality Assurance Team performed a review of the general radiology images and films **[Event 3.22: further investigation]**, while the review of clinical notes, cytology reports and histopathology reports lasted three months (up to December 2005) **[Event 3.23: further investigation]**, and the review of the ultrasound patients' notes and images lasted four months (ending in January 2006) **[Event 3.24: further investigation].** Finally, the review of the 28 patients with delayed diagnosis was performed over December 2005 **[Event 3.25: further investigation].**

5.3.3 Further analysis

Figure 5.18 illustrated three distinct sequences of events that were motivated by detection in different locations.

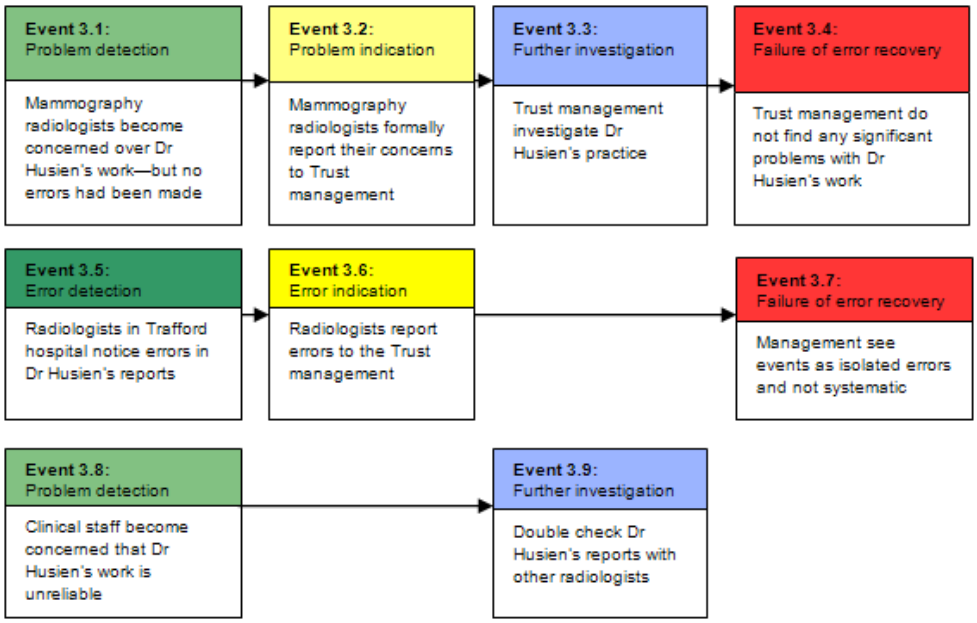


Figure 5.18: Error recovery focused view, Greater Manchester efforts 1.

In the first case, colleagues of Dr H's were concerned about the brevity and the brusque style of his reports. They officially reported their concerns to the Trust management who decided to look into Dr H's past. They found nothing notable apart from a good resume with good references, and assumed that there were no problems with his practice. In the second incident, radiologists in one of the hospitals of the Trust found errors in his reports—this was reported also to the Trust management, but they thought these were isolated events.

However in the third case, it was clinical staff that were concerned and not radiologists. They did not report their concerns—they double checked Dr H's reports with other radiologists. This is common practice in radiology and should have been carried out any way, according to the inquiry report.

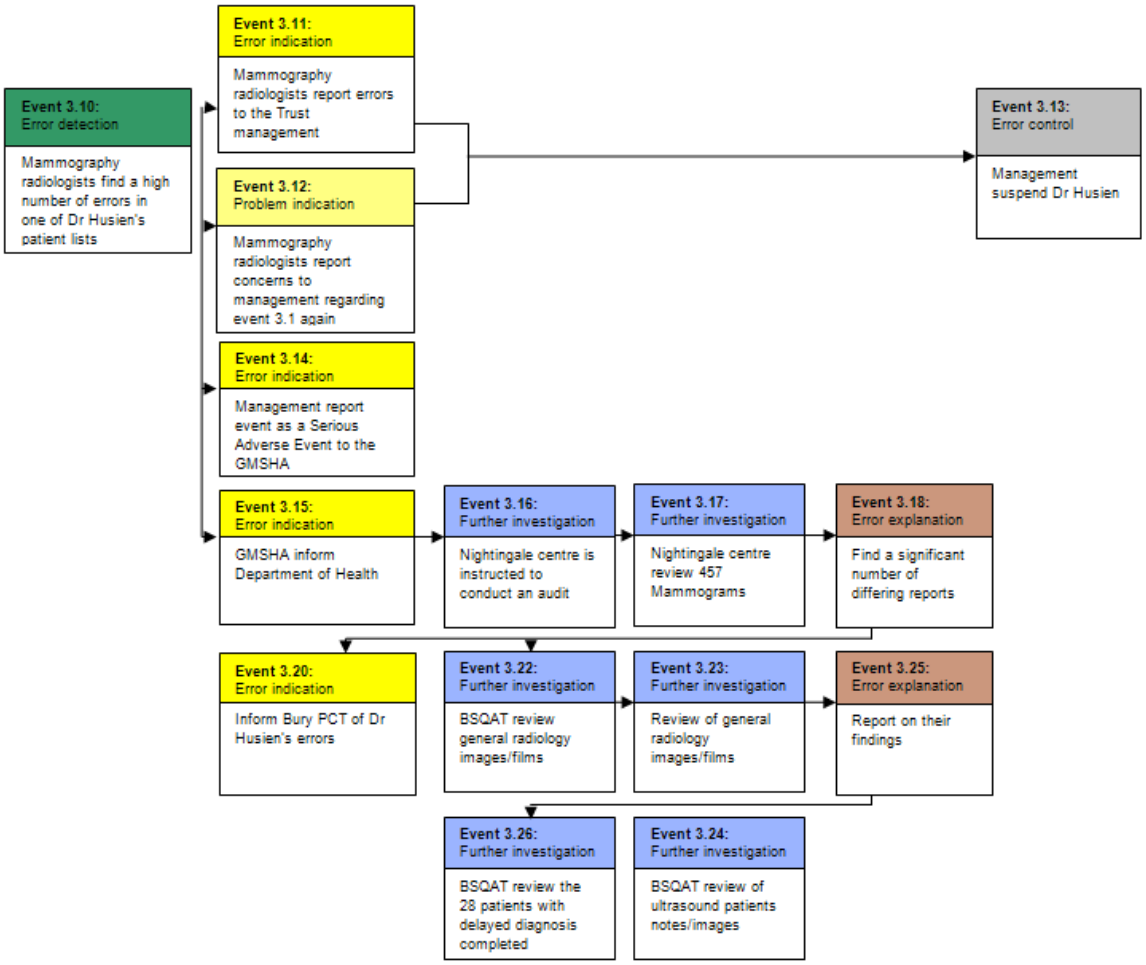


Figure 5.19: Error recovery focused view, Greater Manchester 2.

The incident was directed towards resolution when colleagues of Dr H's found a number of errors in one of his patient lists. Unlike the previous instances of detection, this time reporting of this event had immediate impact. Trust management suspended Dr H.

Investigation activities, like in the previous Breast screening incident, involved multiple inquires—both internal and independent. In a similar manner to the Hammersmith incident, many indication events involved the notification of agencies and authorities in order to establish the inquiries that were necessary.

Event	Analysis
Problem detection	**Outcome based, inside the lab** Concerns over the brevity and overall quality of Dr H's reports by colleagues (event 3.1) **Outcome based, outside the lab** Clinical staff become concerned over the reliability of Dr H's work (event 3.8)
Error detection	**Outcome based, outside the lab** None **Outcome based, inside the lab** Radiologists notice errors in Dr H's reports (event 3.5) Mammography radiologists find a high number of errors in a single patient list of Dr H's (3.10)
Problem indication	Formal reporting of concerns to the Trust management (events 3.2 and 3.12) [presumably in management meetings]
Error indication	Formal reporting of errors to the Trust management (events 3.6 and 3.11) Reporting of incident to the an external authority as a 'serious incident' (regional Health Authority and Department of Health) (events 3.14 and 3.15) Inform other Trust that the liable person works in regarding his performance (event 3.20)
Further investigation	**Inside the lab** To investigate if there is an error (events 3.3 and 3.9) To determine the extent of the error (events 3.16, 3.17, 3.22, 3.23, 3.26 and 3.27)
Error explanation	Identification of reports that contain errors (event 3.18) Compilation of report containing findings (event 3.25)
Error control	Suspend person responsible (event 3.13)
Error correction	n/a (there is no information available regarding subsequent activities to correct the problem)
Failure of error recovery	Investigation does not find any problems (event 3.4) Do not consider errors as systematic, but only as isolated events (event 3.7)

Table 5.9: Analysis of error recovery activities, Manchester 1, 2 and 3.

5.4 Case study 4: STI screening errors, Florida

This last incident presents some similarities with the Sheffield incident due to the involvement of software. However, the bug was not part of the risk calculation but in the reporting scheme; the system which flagged positives and copied them to the test report to be distributed back to where the test was initially requested. In that sense, it also presents some similarities with the second incident, the errors that occurred in the London Breast Cancer Screening service where a confusing notation for denoting positives resulted in missing several positives.

This incident occurred in the USA, where the delivery system is somewhat different from the UK NHS. In the USA there are more analytic laboratories per clinic or surgery, as the Primary/Secondary care distinction does not exist. This had some impact on the resolution of the incident as it was detected within the same hospital and the error had not propagated across various locations.

There is no formal inquiry report for this case study; the analysis was based on an entry in the Risk Digest [Wears, 2004] and further data collected through interviews with involved personnel.

5.4.1 Incident summary

Due to a software bug compounded with interface deficiencies in the Microbiology Department of a hospital in Florida, 275 positive results for Sexually Transmitted Diseases (STDs) were missed over a period of four months in 2003. Consequently, 125 of these cases had not been treated presumptively with antibiotics. Prior to the incident there

was a change in equipment for analyzing DNA probes in the Biochemistry Department. Also, there was a change in the reporting format after a clinician's request which contributed to missing critical test results. During the change of equipment, the ED nurse who was the designated recipient of all tests was away on vacation.

A few years later, a similar error occurred, and the nurse responsible for collecting test results for the Emergency Department, recalled the previous incident and was alarmed very quickly. She contacted the lab shortly after the error (within a week) which found and dealt with without any patient complications. This second incident describes an effective error recovery and not a long- term diagnostic failure as the first one. However, the experience of the first incident contributed greatly to the resolution of the second. For this reason the ER-STEP analysis will cover both incidents.

Incident	Errors in STI Screening
Incident timeframe	February 2003 – June 2003
Primary cause	Software bug
Data source	*Interviews with involved personnel*

Table 5.10: Overview of incident 4.

5.4.2 Background

This incident involved errors in Sexually Transmitted Infections (STI) screening that occurred in a Biochemistry Department, affecting patients in the Emergency Department (ED) of the same hospital in Florida. Although this took place in the USA and not the UK, there are several similarities, as well as significant differences that are worth discussing in comparison to the NHS delivery model.

The software bug was in the system used for reporting results, and not in the actual diagnostic calculations. On a daily basis, the ED would request several different tests to be performed from the Biochemistry laboratory. Once these tests were performed, the reporting system in the Biochemistry Department would only select the positives and not the negatives—these positives are printed and sent to the ED; this was a new system. While the software bug in the reporting system resulted in missing positives, the fact that negative results were not printed as well did not allow for the error being detected.

As mentioned in Chapter 4, data regarding this incident was gathered with interviews with the ED nurse responsible for requesting and collecting reports from the ED, the Head of the Emergency Department and the supervisor of the Biochemistry Department[14].

5.4.3 ER-STEP analysis

Towards the end of January 2003, there was a change in equipment for analyzing DNA probes in the Biochemistry Department. The system for ensuring that results were not missed was based on a custom report written locally. This report covered all bacteriology cultures, not just those for STDs, and looked at a binary field for a positive or negative value for Gonorrhea/Chlamydia (GC) reporting only the positives. Under the new system, that field was empty, and so no GC cases positive were listed by the report **[Event 4.1]**.

[14] This was possible through the 'Ken Browning Traveling Scholarship in Computing and Medicine' of the Department of Computing Science, University of Glasgow.

Because the other cultures were still being reported normally, the report looked normal (ie, it was not entirely empty). Some time after her return **[Event 4.2: no classification]**, the nurse became concerned that there were no positives getting reported **[Event 4.3: problem detection]**. She discussed the matter with her supervisor **[Event 4.4: further investigation]**, who advised her to report this to the Biochemistry Department **[Event 4.5: further investigation]**; she shortly after visited the lab and informed staff **[Event 4.6: problem indication]**.

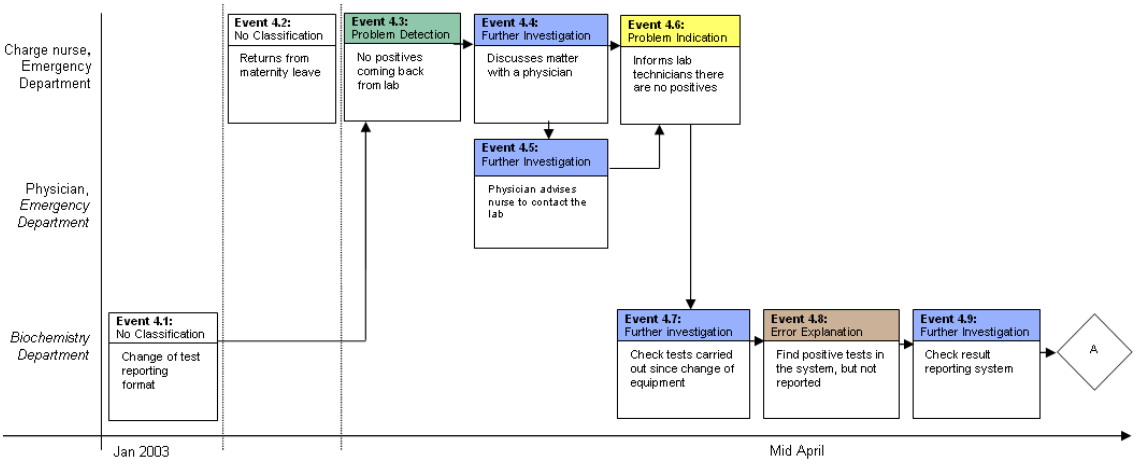


Figure 5.20: Error recovery activities 1, Florida.

Technicians looked at tests dating back to the time of the installation of new equipment **[Event 4.7: further investigation],** and found positives had been diagnosed, but were not printed in the report sent to the ED **[Event 4.8: Error explanation]**. They then checked the reporting system **[Event 4.9: further investigation]** and identified the bug **[Event 4.10: Error explanation]**. The custom program was amended **[Event 4.11: error correction, Event 4.12 error correction]**. All involved patients were identified **[Event 4.13: further investigation]** and patients were contacted [**Event 4.14 error correction**].

In cases of STI screening errors, the patients had to be asked to inform their recent sexual partners to be screened as well.

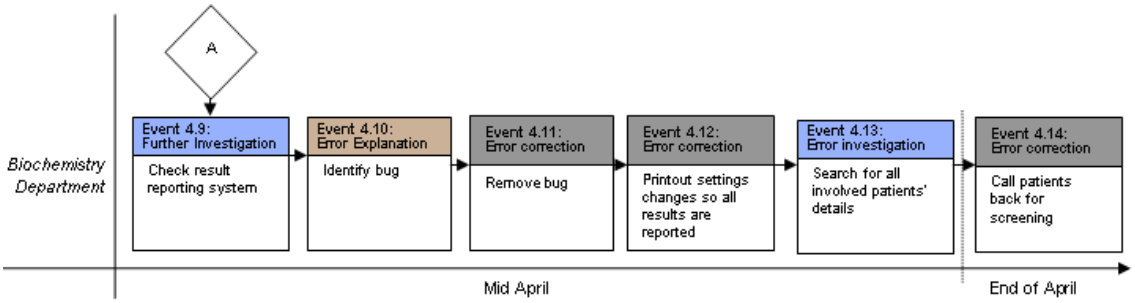


Figure 5.21: Error recovery activities 2, Florida.

Two years later, the same ED nurse noticed after 2 weeks that no positives were coming back **[Event 4.15: problem detection]**. According to her experience, there were 2 or 3 positives every week. As she recalled the previous incident **[Event 4.16: No classification]**, she immediately visited the laboratory **[Event 4.17: problem indication]**, and after checking the system **[Event 4.18: further investigation]** it was discovered that a 'flag' used to check whether a test result is positive had been mistakenly deactivated **[Event 4.19: error explanation].**

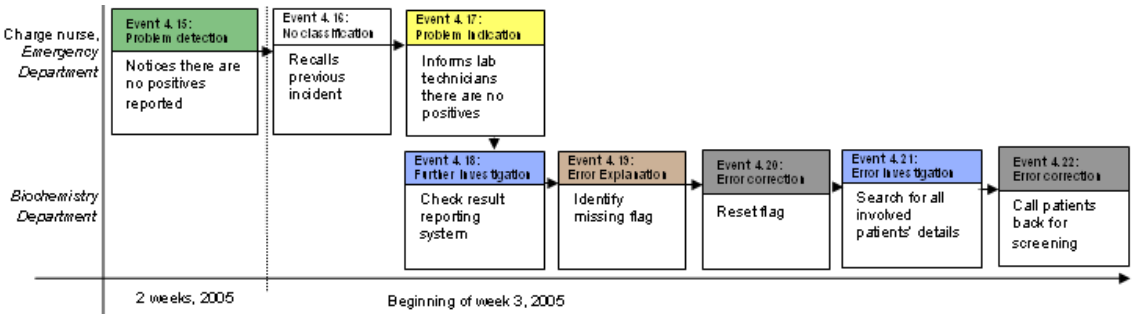


Figure 5.22: Error recovery activities 3, Florida.

The flag was reset **[Event 4.20: error correction]** and the system continued to be used without any further problems. All involved patients were identified **[Event 4.21: further investigation]** and patients were contacted [**Event 4.22 error correction**]. The incident did not last long enough to have any adverse impact on patients' health.

5.4.4 Further analysis

Figure 5.23 illustrates the error recovery focused view of the ER-STEP analysis presented in Figure 5.20. As we can see, there was only one instance of problem detection which was however enough for an effective recovery process to be carried out. This is perhaps due to the fact that problem detection occurred in the same hospital as the laboratory. There are also some significant differences in the model of diagnostic services delivery between the UK and US (as discussed in section 2.2) which perhaps play in role in the US incident response requiring less effort from the person who detected the problem/error. This will be discussed in more detail in the next chapter.

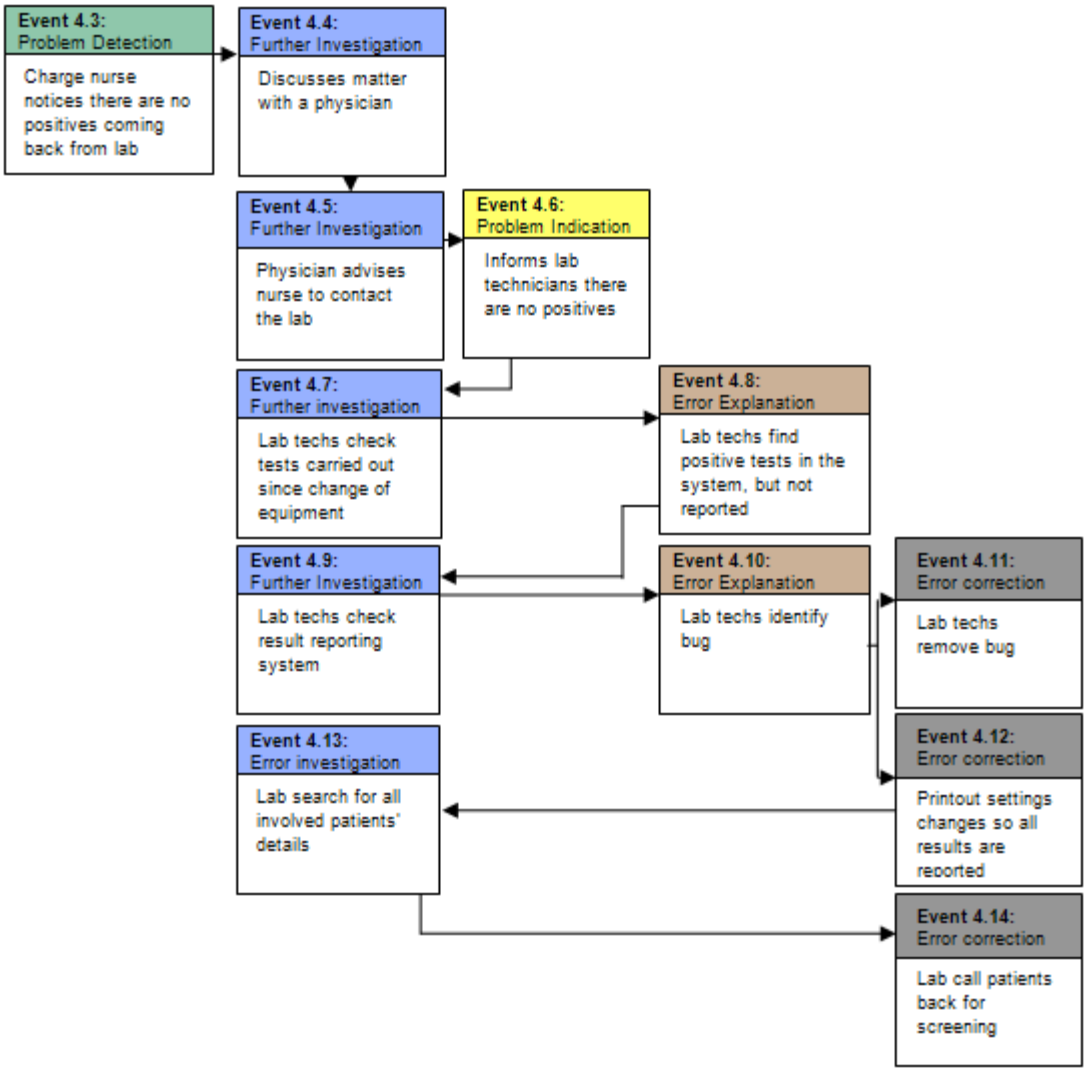


Figure 5.23: Error recovery focused view 1 and 2, Florida.

In a similar manner to the previous incidents, the timeframe of the Florida incident is extended by the investigation required to identify and contact all involved patients for re-screening. This problem is of particular significance when considering infectious diseases as more people may be affected over time.

Figure 5.24 illustrates the error recovery focused view of the second incident that occurred in the same hospital, involving the same ED nurse as the previous occurrence.

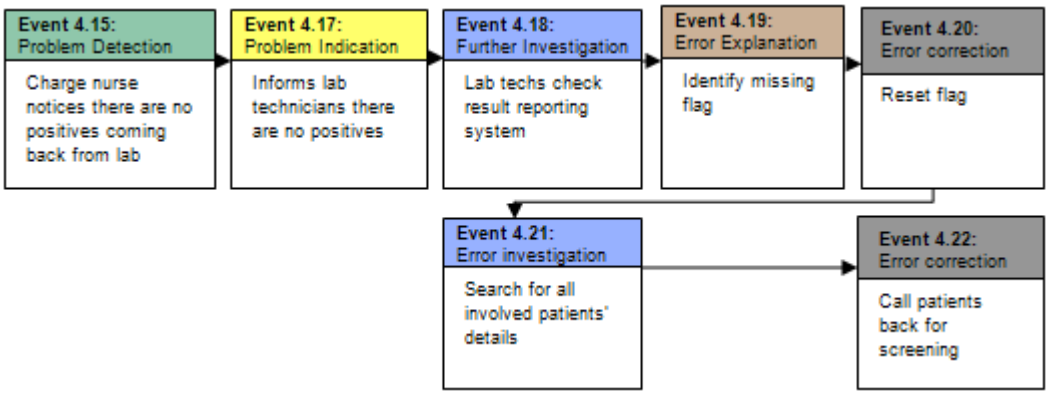


Figure 5.24: Error recovery focused view 2, Florida.

In this case, the experience of the previous incident helped in making a quicker and more efficient recovery. This incident only lasted two weeks and there were no affected patients from the error.

Event	Analysis
Problem detection	**Outcome based, inside the lab** none **Outcome based, outside the lab** ED charge nurse becomes concerned over the lack of positives reported from the lab (events 4.3 and 4.15)
Error detection	**Outcome based, outside the lab** None **Outcome based, inside the lab** none
Problem indication	Nurse reports concerns to the lab (face-to-face communication) (Events 4.6 and 4.17)
Error indication	None
Further investigation	**Outside the lab** To decide what action to take (events 4.4 and 4.5) **Inside the lab** To investigate if there is an error (events 4.7 and 4.18) To determine the extent of the error (events 4.13 and 4.21)
Error explanation	Identification of erroneous reports(events 4.8)

Event	Analysis
	Identification of technical problem (events 4.10 and 4.19)
Error control	None
Error correction	Fix technical problem (events 4.11, 4.12 and 4.20) Call involve patients and their partners back for screening (events 4.14 and 4.22)

Table 5.11: Analysis of error recovery activities, Florida.

5.5 Chapter summary

This chapter presented the ER-STEP analysis of four screening incidents which were severely prolonged by a late detection and poor incident response. The analysis with ER-STEP diagrams, the restructured view and a categorization and summary of all types of activities that took place according to the error recovery framework stages has been useful in understanding *what* happened in these incidents in terms of error handling.

The purpose of the next chapter is to further analyze these incidents, by integrating and comparing them, in order to draw high level conclusions about the factors that inhibit effective detection and recovery in screening programmes.

Chapter 6: Overview of findings

Having analyzed the four incidents with the same technique, we can now systematically compare and integrate the individual analyses' findings and draw high level conclusions about key problem areas that may impede an effective error recovery in screening services (with perhaps implications for other laboratory services as well).

The purpose of this chapter is therefore to attain a high-level perspective on screening error handling. The tables produced at the end of each of the four case studies will be used to aggregate and categorize the different kinds of activities that can be seen to fall under a specific stage of the error recovery framework—this will help in gaining an understanding about individual stages. However, it is also important to understand the relationship between the different stages, and the different activities. For instance, what are the different possible activities that may follow outcome based problem detection taking place outside the lab? What are the different kinds of further investigation, and what events may trigger these activities? This chapter will conclude with the 'screening error recovery model'—this is based on the error recovery framework which has been enriched by the findings of the analysis of the four incidents.

6.1 Error recovery stages

6.1.1 Problem detection

Table 6.1 presents the different kinds of problem detection that occurred in the four case studies. As we can see, problem detection is most likely to occur outside the lab (i.e. where the test results are used). In two cases, problem detection outside the lab took the form of growing concerns regarding a change in the frequency of positives/high-risk patients. Also, in all cases of problem detection outside the lab, it was nurses (in one case presumably physicians as well) who started becoming concerned. In addition, the people who experienced problem detection also acted on upon their own initiative to report these concerns to the lab. This illustrates the importance of the nurses responsible for requesting and following up on test results, as they are the ones most likely to start becoming concerned.

	Incident 1: Sheffield	**Incident 2: London**	**Incident 3: Manchester**	**Incident 4: Florida**
Outcome-based problem detection *outside* the lab	*[Events 1.4 and 1.23]* Concern over low number of high-risk pregnancies	—	*[Event 3.8]* Clinical staff have concerns over the reliability of Dr H's work	*[Events 4.3 and 4.15]* Concern over low number of positives
Outcome-based problem detection *inside* the lab	—	—	*[Event 3.1]* Radiologists have concerns over the brevity of Dr H's reports	—

Table 6.1: Problem detection events

Problem detection inside the lab occurred only in the Manchester incident. Colleagues of the person responsible were increasingly becoming concerned as the reports he was

compiling were too short and brusque, while they had realized he was not following all appropriate procedures.

Findings regarding problem detection can therefore be summarized as follows:

- **Problem detection outside the lab (outcome-based):** This kind of detection was performed primarily by the nurse responsible for requesting and collecting test results (**events 1.4, 1.23, 4.13 and 4.23**); the nurse will become increasingly concerned as his/her expectations of the frequency of positives/negatives reported back from the lab drops. Problem detection may take weeks to lead to some further action.
- **Problem detection inside the lab (outcome-based, could be action-based):** This kind of detection only occurred once in the four incidents **(event 3.1)**. This took place when colleagues of Dr H had concerns that his reports were short. No error had occurred though, so this event is labeled as problem detection. Problem detection arising from the evaluation of test results, as occurred outside the lab, did not take place. Therefore, problem detection inside the lab may rarely occur by expert clinicians who realize a procedure is not carried out as prescribed, either by examining a report, or by observing the conduct of the person responsible.

6.1.2 Error detection

Unlike problem detection, *error detection* does not build up over time or come with uncertainty; this is because there is evidence that an error has occurred. Several different kinds of error detection occurred in the four case studies; most of them involved a specific patient rather than a trend in the frequency of positives/negatives.

Apart from the action-based, outcome-based, through limiting functions and location based classification, there was another kind of error detection that was found in the case studies: *detection through further investigation*. Such an occurrence highlights the systemic nature of error recovery where problem detection in one organization may lead to error detection in another. Table 6.2 below presents these various events of error detection that occurred in the four case studies.

Detection type and location	Incident 1: Sheffield	Incident 2: London	Incident 3: Manchester	Incident 4: Florida
Outcome-based error detection *outside* the lab	*[Event 1.6]* Incident with one patient	*[Event 2.4]* Revaluation of a patient's screening result by new breast screening service found error	—	—
Outcome- based detection *inside* the lab	*[Event 1.41]* When trying to change DOB, risk calculation remained the same	—	*[events 3.5 and 3.10]* Identification of errors in Dr H's reports	—
Action- based detection in the lab	—	—	—	—
Error detection through further investigation *inside* the lab	*[Event 1.62]* Accumulation and investigation of audit data results in noticing a lack of high- risk results	—	—	—
Error detection through further investigation *outside* the lab	—	—	—	—

Table 6.2: error detection events.

Findings regarding error detection can be summarized as follows:

- **Outcome-based error detection outside the lab:** Primarily incidents involving specific patients **(events 1.6 and 2.4)**. This type of error detection will occur

when there are obvious errors with the test results, either because they do not make sense at all, or because clinicians have already established a set of potential diagnoses which is obviously contradicted in the test results. In addition, we should take into account accidental detection, as it occurred in the London Hammersmith incident. A cancer patient which had been sent the wrong mammogram been cleared of the diseases. When she moved to another Breast Screening Service, she was diagnosed again, unveiling the error of the previous diagnosis.

- **Outcome-based error detection inside the lab:** Audits and many other quality assurance practices aim to detect such errors before test results leave the lab **(events 1.41, 3.5 and 3.10)**. However, the incidents discussed here have occurred because errors were missed by the lab. Outcome based error detection in the lab considers all other possible ways through which the outcome of laboratory work is evaluated against errors. This kind of detection during an incident is rare, as, it has already been suggested that laboratories maintain little or no information about the subsequent patient outcomes of their work, which limits their evaluation of testing practices (Bonini et al., 2005). It is therefore very difficult for labs to evaluate the validity of test results; something which can take place in points of care where clinicians have physical contact with the patients when they integrate test results in the diagnostic process. Outcome based error detection inside the lab occurred only in the Manchester incident, where radiologists found errors in Dr H's work **(events 3.5 and 3.10)**.

- **Action-based error detection in the lab:** Action-based detection can only occur in the lab, as it refers to error detection during the process of analyzing a specimen and performing the subsequent calculations to compile the patient's test results. However such detection was not identified in the four case studies. This is a kind of detection that should be supported and will be discussed in the recommendations that are put forth in the next chapter.
- **Error detection through further investigation in the lab:** This kind of error detection occurred in the Sheffield incident **(event 1.62)**. Following reports from points of care, investigation in the lab was carried out to find out if and what is actually wrong. Error detection is considered to be the first stage in error recovery [Zapf and Reason, 1994]; however, as we have seen here, if problem detection occurs, error detection may come several stages later. This kind of detection is very much a system function, and involves further investigation and problem indication to take place first.
- **Error detection through further investigation outside the lab:** There were no such instances in the four case studies. This makes sense as investigation activities in points of care were very limited; they could only monitor the trends of test results, consult with colleagues or make an enquiry to the lab; all activities which may lead to problem detection, but not error detection.

6.1.3 Further investigation

Further investigation includes a set of diverse activities, ranging from monitoring of results, review of audit results, communications to investigate aspects of the problem,

meetings, and internal and/or external inquiries. Furthermore, further investigation may take place both in and outside the lab.

In order to proceed to a useful classification of all further investigation activities they were categorised not only according to location, but also according to purpose. The following purposes could be identified:

- To determine if there is an error
- To determine what the error is and its causes
- To determine the extent of the failure

Table 6.3 summarizes and categorizes all the further investigation events that were identified in the four case studies.

Investigation type and location	Incident 1: Sheffield	Incident 2: London	Incident 3: Manchester	Incident 4: Florida
To decide what action to take, *outside* the lab	*[Event 1.32]* Enquiry to colleague	—	—	*[event 4.4]* *Enquiry to colleague*
To decide what action to take, *in* the lab	*[Events 1.19 and 1.35]* Enquiry to colleague	*[Event 2.18]* Meeting to decide further action	—	—
To find out if there is an error, outside the lab	*[Event 1.12]* Monitoring of results	—	—	—
To find out if there is an error, in the lab	*[Events 1.38, 1.39, 1.55, 1.61 and 1.63]* Arrangement and conducting of audit	*[Events 2.22 and 2.25]* Arrangement and conducting of audit	*[Event 3.3]* Investigate Dr's past and background *[Event 3.9]* Double check Dr's reports	*[events 4.7 and 4.18]* Investigation of system

Investigation type and location	Incident 1: Sheffield	Incident 2: London	Incident 3: Manchester	Incident 4: Florida
Enquiry regarding past activities, outside the lab	—	—	—	—
Enquiry regarding past activities, in the lab	*[Event 1.36, 1.54 and 1.57]* Discussing whether an audit has been carried out or not	—	—	—
To find out what the error is, outside the lab	—	—	—	—
To find out what the error is, in the lab	*[Event 1.66]* Investigation of system	*[Events 2.19, 2.20, 2.28 and 2.30* Investigation of system	—	—
To determine the extent of the error, outside the lab	—	—	—	—
To determine the extent of the error, in the lab	[event 1.70] Examination of all test results since Jan 1st	*[Events 2.22 and 2.25]* Audit	*[Events 3.16, 3.17, 3.22, 3.23, 3.26 and 3.27]* Audit	*[events 4.13 and 4.21]* Checking of all test results

Table 6.3: Further investigation events.

Findings regarding further investigation are therefore as follows:

- **Further investigation outside the lab, to decide what action to take:** This is most likely to occur following problem detection. The nurse who performed problem detection may seek advice as to what action to take. In the Sheffield incident, this took place after some attempts of problem indication had failed **(event 1.32)**, while in the Florida incident this occurred straight after problem

detection **(event 4.4)**. In addition, the colleague that gives the advice may be a supervisor or a physician.

- **Further investigation inside the lab, to decide what action to take:** This kind of investigation will take place when people in the lab have acknowledged the possibility that something might be wrong. This may be informal (e.g., asking a colleague for advise) **(events 1.19 and 1.35)** or formal (Department, hospital or event Trust level meeting) **(event 2.18)**.
- **Further investigation outside the lab, to find out if there is an error:** In the four case studies, there was only one such occurrence: In Sheffield, a nurse that had performed problem detection was monitoring her own results **(event 1.12)**. As there are very little means for staff outside the lab to investigate, this kind of further investigation is fairly limited as to what it may achieve.
- **Further investigation to find out if there is an error, inside the lab:** This kind of investigation is the one most likely to have a major impact on the success of an error recovery process, as it may directly lead to *error explanation*. When it is in the form of an audit, it may take a significant amount of time, but this can also lead to—at least—an initial estimate of the number of patients misdiagnosed. Audits were requested and carried out in the Sheffield and London incidents **(events 1.38, 1.39, 1.55, 1.61 and 1.63, and 2.2, 2.5)**. An investigation into the system, whether software or organizational process, will also fall under this category of further investigation.
- **Further investigation to find out what the error is, inside the lab:** There may be an overlap between this kind of investigation and investigation to find out if

there is an error as an investigation to establish the presence of an error will be based on an informed hypothesis of the operator or other involved actor. However, in case where further investigation to find out if there is an error was in the form of an audit of test results, more investigation in the form of examination of the system (software/hardware, previous test results) will be required in order to establish what the error is.

- **Enquiry regarding past activities, inside the lab:** It was found that in many cases, audits regularly requested had not been carried out. This emerged when in the inquiry report the person who had requested the audit asked weeks, or months later if that audit had been carried out or not. Although that audit may not form part of the error recovery if it was a routine activity, the enquiry regarding whether it has been carried out or not may well do, if it has been triggered by recovery related activities (e.g., by problem indication) **(events 1.36, 1.54 and 1.57**).
- **To determine the extent of the error, inside the lab:** This may take place either after or during the final stages towards the correction of the causal factors that led to the errors in the first place. In a screening incident, this is a very important part of error recovery as errors in screening will continue to have an impact as time passes on the people misdiagnosed. In the STI incident that took place in Florida, there were serious social implications as infected patients who were told they are free from disease could have possibly infected others which could remain unknown. Diagnostic services must therefore be prepared to contact involved

patients as soon as possible. Such kind of investigation may also involve independent auditors, which would severely prolong the incident timeframe.

6.1.4 Problem indication

Problem and error indication are very crucial—it was found that indication events were ones most likely to fail to achieve a progression to a later error recovery stage. Problem and error indication are most likely to be initiated outside the lab, following problem and error indication. It was observed that various means of communication were employed; these are used to classify the indication events that took place.

Table 6.4 summarizes and categorizes the different kinds of problem and error indication events that took place in the four incidents, taking into account the various means of communication that were used.

Problem Indication	Incident 1: Sheffield	Incident 2: London	Incident 3: Manchester	Incident 4: Florida
Face-to-face	—	—	—	*Events 4.6 and 4.17]* ED charge nurse visits Microbiology lab and informs them of the absence of positives
Over the phone	*[Event 1.3, 1.5, 1.7, 1.24, 1.26 and 1.34]* Nurse that performed problem detection phones the lab	—	—	—
Written	*[Event 1.34]* Nurse sends formal letter (fax)	—	—	—

Problem Indication	Incident 1: Sheffield	Incident 2: London	Incident 3: Manchester	Incident 4: Florida
Formal/ incident reporting	—	—	*[event 3.2 and 3.12]* To Trust management in management meetings	—

Table 6.4: Problem indication events.

The following conclusions have been reached regarding problem indication events and means of communication used:

- **Problem Indication over the phone:** As problem detection most commonly occurred in primary care institutions, it is understandable that the use of the phone for the purposes of problem indication was the most frequently used means of communication (6 occurrences in the Sheffield case). However, it was found to be greatly ineffective. Reasons for the failure of problem indication were two: the lack of evidence during the claim, and the breakdown of communication; two factors which are also intertwined and will be discussed later on in this thesis.
- **Written problem indication** occurred only once in the four incidents **(event 1.34)**. A letter was directed from the hospital which detected the problem to the lab. It should be noted that no incident reports were written, even though incident reporting schemes where present in at least the three NHS incidents.

6.1.5 Error indication

Table 6.5 summarizes and categorizes all types of error indication events that took place in the four incidents in a similar manner with problem indication events discussed in the previous section.

Indication	Incident 1: Sheffield	Incident 2: London	Incident 3: Manchester	Incident 4: Florida
Face-to-face	—	—	—	—
Over the phone	*[event 1.45, 1.46 and 1.48]*	*[events 2.5, 2.9 and 2.12]*	*[event 3.20]* Trust management inform other hospital of Dr's errors	—
Voice message	*[event 1.65]* Leaves voicemail after person is not available on phone	—	—	—
Written	—	*[event 2.6 and 2.10]*	—	—
Formal/ incident reporting	—	*[events 2.14, 2.15 and 2.16]* Reporting to regional or national authorities	*[events 3.6 and 3.11]* To Trust management *[events 3.14 and 3.15]* Regional and national health authorities.	

Table 6.5: error indication events.

The following conclusions can be drawn about error indication:

- **Face-to-face error indication** did not occur in the four incidents. This kind of detection is limited due to the physical separation between points of care and laboratories in the NHS. As error detection is most likely to occur where test results are used, face-to-face error indication is going to take place in the cases where errors are detected within the same hospital.
- **Error indication over the phone:** Like problem indication, error indication was mostly done over the phone, as the errors were detected outside the physical

premises of the lab and host hospital **(event 1.45, 1.46, 1.48, 2.5, 2.9, 2.12 and 3.20)**.

- **Written error indication:** Letters were written to the lab and to national agencies when the error was deemed to be of great importance. However this was done as a last resort and only when it when initial reports to the potentially responsible organization were not addressed **(event 2.6 and 2.10).**
- **Formal/incident reporting:** Incident reporting schemes that were in place within hospitals and diagnostic services were in fact not used. This involved cases where the management of a health organization decided to formally inform an authority of significant errors of other organizations. Informed agencies included regional Quality Assurance Centers, the National Patient Safety Agency and the Department of Health. Formal reporting to national agencies occurred in the London and Manchester incidents.

Digital communication, such as email were not mentioned in the four case studies. There was one instance of voicemail, which was disregarded by the recipient; however the caller called again the next day, and the recipient was reached.

Communication breakdowns will be discussed in a separate section in this chapter; this is because there are some common issues with other stages, and primarily further investigation.

6.1.6 Error explanation

Error explanation is ultimately a result of a successful investigation into what the error is, and what the extent of the failure is. In most cases, reaching error explanation meant that error control and recovery are feasible at that stage.

6.1.7 Error control and recovery

Error control took the form of suspension of the person responsible in the Manchester incident **(event 3.13)** and suspension of breast screening until the investigation was completed in the London incident **(event 2.29).** Error control will establish certainty that no further errors will take place. Like error explanation, there is little to add to this error recovery stage, as the problematic areas are the ones earlier on in the recovery process.

6.1.8 Failure of error recovery

These are events where the progression from one error recovery stage to the next has been halted. These are primarily communication breakdowns (see Table 6.6 below).

Failure of error recovery	Sheffield	London	Manchester	Florida
Various activities	*[event 1.11]* Incident not logged *[event 1.10]* Reassurances that 'everything is OK' *[events 1.25, 1.28 and .1.51]* Previous events not considered as important	*[event 2.8]* Previous events not considered as important	*[event 3.4]* Investigation does not find any problems *[event 3.7]* Do not consider errors as systematic, but only as isolated events	—

Table 6.6: Failure of error recovery events.

Failure of error recovery suggests the failure of the previous type of event. Primarily, these will either be problem/error indication or further investigation. Failure of error recovery will either result to the reiteration of previous activities, or to bringing the entire recovery process to a halt.

It is notable that all of these failure of error recovery events have taken place after problem/error indication has occurred. In other words, the actual failure to progress towards error explanation and control/recovery involved people who were expected to take mitigative action. So, they either dismissed a report by the people who performed detection, or they conducted a poor investigation.

Ignored indication events can be seen as communication breakdowns, as the people who are in communication fail to achieve *shared understanding* [Dix et al., 2004]. The

reasons for this are part of the discussion of further analysis of communication events of problem/error indication and further investigation (section 6.3).

The next section will present an overview of the findings that have been discussed so far in this chapter. This is a particularly important part of this thesis, as it presents a model that considered all of the identified sets of activities that fall under each of the stages of error recovery, and their relationship towards the achievement (or failure of) error recovery.

6.2 The screening error recovery model

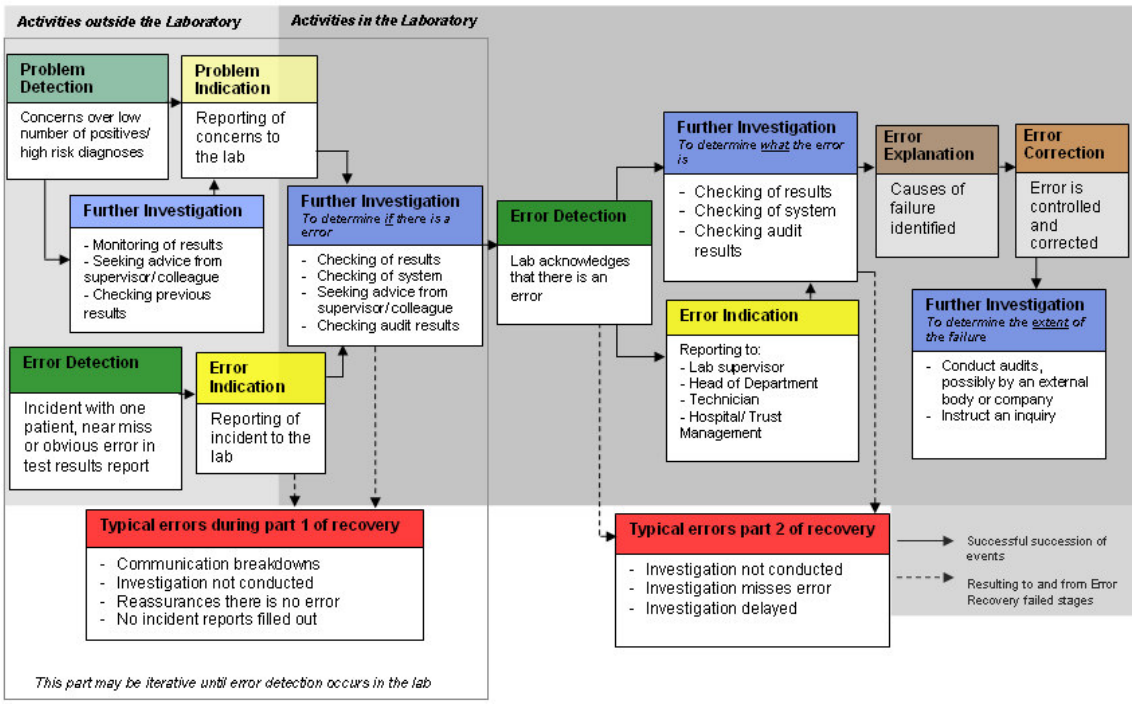


Figure 6.1 illustrates the *screening error recovery model*. This model is based on the error recovery framework, which has been enriched by the findings of the analyses conducted in the past two chapters.

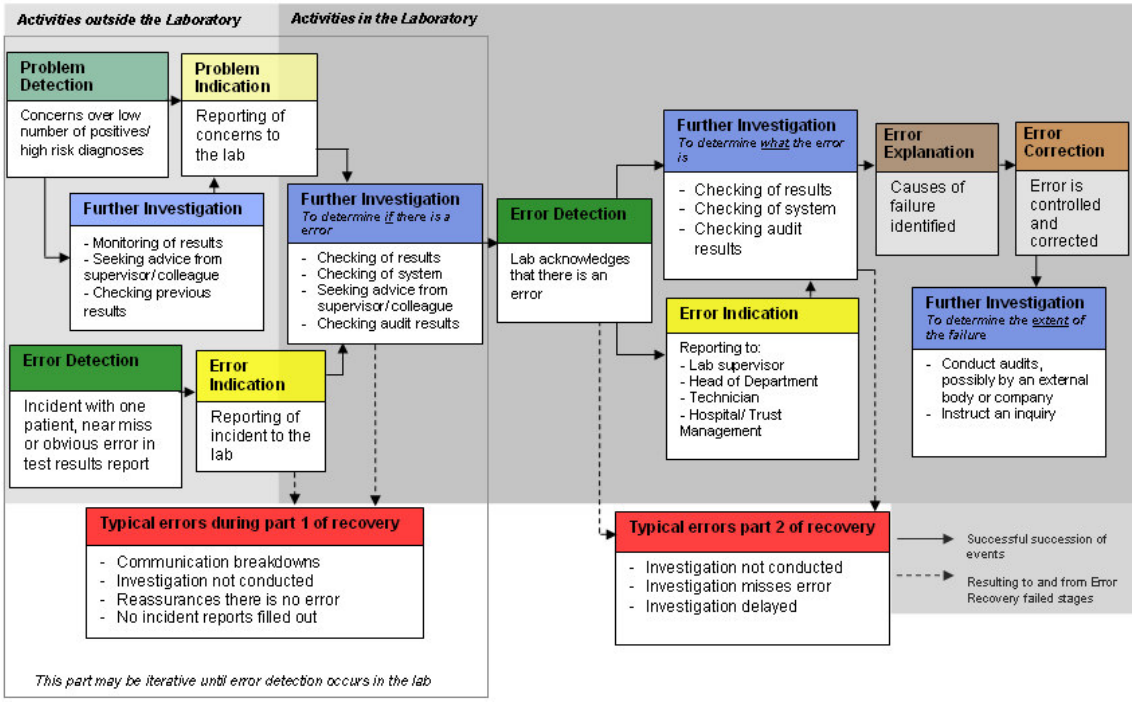


Figure 6.1: The Screening error recovery model.

The lighter shade of gray indicates the set of activities that are possible outside the lab (i.e. where test results are used), while the darker shade of grey illustrates the set of activities that may take place within the lab.

The process can be seen in two parts:

- **Part 1:** *Until error is acknowledged in the laboratory*. Part 1 includes the initial activities that will start where test results are requested and used. Most likely to be motivated by problem detection, the nurse responsible for request and collection of test results will either report his/her concerns to the lab, monitor his/her own results, or report his/her concerns directly to the lab.
- **Part 2:** *Identification and correction of errors.* Having acknowledged the presence of an error, laboratory technicians will investigate to determine what the nature of the error is and what has caused it. This may take time as an audit might have to be called for. Communication of the incident will either be internal

(within the Department, hospital or Trust) or it may involve NHS authorities such as regional quality assurance centers, the National Patient Safety Agency and the Department of Health. The recovery process may be prolonged even if an independent auditor is required to examine large volumes of test results in order to determine the extent of the failure (i.e. the number of patients that have been misdiagnosed). In the case studies analyzed here, the acknowledgement of an error within the lab was not always clearly noted, but occurred along with error explanation following an investigation into claims coming from outside the lab.

An advantage of this model is that it illustrates at what stage different sub-types of error recovery stages may occur (e.g., further investigation to find out if there is an error etc). It can also be used to identify and relate activities that will take place in the different parts of the healthcare system. Eventually, the purpose of this model is to improve individual error recovery activities and consider the subsequent communication link between points of care and the laboratory. However, it is still abstract at this stage and should be enriched with particular information that pertain to a specific screening process, fitted within practices and regulation of a real diagnostic network. This model could be seen as a starting point for considering the design of error recovery strategies customized according to a particular setting.

The following section will discuss the key conclusions that arise from this analysis.

6.3 Key problem areas

It is important to stress that the *failure of error recovery* is not the only reason why the timeframe of an incident may be prolonged. Failure of failure recovery will interrupt the process, but other problems may slow it down. The recovery processes we have seen were mostly initiated by problem detection, and not error detection. Problem detection however may take a significant amount of time to occur. In addition, the lack of evidence that comes along with problem detection is very likely to compromise the effectiveness of problem indication. *Limited detection*, along with the fairly loosely defined process for reporting concerns over the phone and, on the other end, dealing with complaints lead to extended failures to detect and recover from errors (*communication breakdowns*).

These two main factors (limited detection and communication breakdowns) will be briefly discussed here, and will be fully analyzed in the next chapter, which will also consider implications for design, and recommendations for the improvement of laboratory error handling.

Limited detection

Limited detection can be seen to be due to a **lack of appropriate feedback** [Norman, 1983] from the system. The lack of appropriate feedback did not allow for error detection, and especially for action-based detection, which is an immediate way of identifying errors. Detection had to therefore take the form of problem detection, and was based on the experience of the nurse responsible for requesting and receiving test results. However, problem detection and the lack of appropriate feedback are not only

responsible for a late detection, but also a determinant factor of the failure of *problem/error indication* and *further investigation*.

The lack of evidence often resulted to **communication breakdowns**, especially in inter-organizational settings, which prolonged the incident adversely.

Communication breakdowns

Problem and *error indication* are the reporting of concerns of detected errors to the laboratory and other key parties, expecting them to act upon these reports. This should not be confused with communication during *further investigation*, where the person who initiates the communication is not reporting an error, but is trying to find out what caused it. However, common problems may affect both.

Communication failures are an important contributor to adverse events in medicine. In a review of 14,000 in-hospital deaths, communication errors were found to be the lead cause [Wilson et al., 1995], while about 50% of adverse events detected in a study of primary care physicians were related to communication difficulties [Bhasale et al., 1998].

Breakdowns during conversation are relatively frequent occurrences; however we tend to be able to repair them when we communicate [Winograd and Flores, 1986]. Breakdowns occur due to divergence in topic focus, due to ambiguity in a speaker's expression, or merely because someone misheard a word. Most breakdowns are detected quickly, but in

many cases people might think they have achieved a *shared understanding*, while in fact, they have not (Dix et al., 2004).

Many of these breakdowns occurred because reporting was informal. In the Sheffield incident, several reports over the phone were ignored, but a single letter stimulated some activity to investigate if there was actually a problem [see **event 1.34**, p. 89]. This highlights the danger associated with informal reporting; however, if formal reporting is introduced without appropriate consideration, there is a danger that it creates barriers to reporting; people may be more reluctant than mentioning something to a colleague.

There is much work currently done in understanding communication breakdowns in healthcare. For instance, communication breakdowns in the operating room [e.g. Lingard et al., 2004 or Greenberg et al, 2007], or during patient hand-offs [e.g. Solet et al., 2005 and Patterson et al, 2004]. Such work could be considered to further analyze these instances of communication breakdown that were found in this study.

The purpose of the next chapter is to address the issues of limited feedback and communication breakdowns further, and to generate a set of useful recommendations that could be used to improve these two aspects of error handling. These recommendations are also based within the error recovery framework and the screening error recovery model, and will be related to the different stages of the recovery process.

6.4 Chapter summary and discussion

This chapter presented the final stage of analysis of this thesis. The findings of the four individual incident analyses were aggregated, compared and integrated in order to understand how errors in screening programmes are detected and handled. This analysis resulted to the screening error recovery model, which summarizes the various activities that fall under each of the recovery stages, as well as their relationship. This model can be used to devise error recovery strategies for screening programmes.

The chapter concluded with the identification of two general problem areas: the lack of appropriate feedback and communication breakdowns. These two problem areas affected various stages of error recovery in different ways. Having abstracted to a relatively high-level, we can now relate these two problem areas to the various stages, and consider how to deal with the particular problems. The next chapter will discuss these two issues in more detail, and suggest a number of recommendations that may be applicable to screening services and diagnostic services in general for a more effective detection and recovery of errors.

Chapter 7: Recommendations

The recommendations that will be discussed in this chapter address the challenges that have been discussed in the previous chapter. In order to produce recommendations on the improvement of screening error handling, the screening error recovery model is used in combination with systems design principles. In particular, they draw upon theory presented in *Chapter 3: Theoretical context*, theory in Human Computer Interaction (HCI), but also focus on the Quality Assurance practices and related artefacts that were discussed in *Chapter 2: Field and focus* and during the analysis of the case studies (such as auditing and incident reporting). These recommendations consider the improvement of individual stages of the screening error recovery model, but also take into account their impact on the subsequent stages, e.g. how a particular improvement in *problem detection* may facilitate better *problem indication*.

There are two important limitations of the recommendations put forward in this chapter: first of all, they are *high-level*, as they have been abstracted from the laboratory context of the individual incidents that have been analyzed; this was necessary in order to reach a level of generalization that would cover the diverse set of screening services available. Therefore, their application will require further analysis and instantiation so that they can be focused on the particular processes, job roles, technology and regulation.

A second limitation is that these recommendations *lack validation*. Some of these have formed part of publications produced during this work [Chozos, 2008; Wears et al., 2008], but this does not stand as sufficient validation for their application in an actual medical context at this stage. It is also likely that some of the suggested practices are also currently existing in NHS diagnostic services; this is because the recommendations are based on the findings of the analysis of the four incidents but a thorough review of actual systems has not been conducted.

Nonetheless, these recommendations are important as they expand on the findings of the analysis of the four incidents and could suggest general principles for good practice and directions for further research. The discussion on the relationship between the different stages should be considered as the most significant contribution of this chapter.

Table 7.1 illustrates the four different kinds of recommendations that are put forward in this chapter, and their area of relevance.

	Checking patterns	Interface design	Software app	Communication
Problem detection	—	—	Section 7.1.3	—
Error detection	Section 7.1.1	Section 7.1.2	—	—
Action-based	Section 7.1.1	—	—	—
Outcome-based	Section 7.1.1	Section 7.1.2	—	—
Limiting functions	—	Section 7.1.2	—	—
Problem indication	—	—	Section 7.1.3	Section 7.2
Error indication	—	—	Section 7.1.3	Section 7.2
Further investigation	—	—	Section 7.1.3	Section 7.2

Table 7.1: Recommendations and relevance to error recovery stages.

7.1 Detection

Recommendations for improving detection will consider:

- Checking patterns
- User interface issues
- Software applications

7.1.1 Checking patterns

Action-based detection (discussed in Section 3.3) would occur during the analysis of a specimen or while entering data in a computer system. This primarily focuses on instances of human error, and not systematic errors that have been the focus of this thesis. Furthermore, action-based detection is very much dependent on the nature of the task which varies in relation to the diagnostic service—specific guidelines applicable here therefore should be driven by detailed insight from laboratory medicine.

Human redundancy (discussed in Section 3.4.2) [Clarke, 2005] is applicable here as it can help to improve detection of errors in the lab during or shortly after the analysis; the concept is based on the presence of a colleague observing the person who is performing a specimen analysis. Human redundancy suggests the following:

1. **One person checks the outcome of their colleague's work.** In a laboratory setting, a test result is evaluated by another colleague.
2. **A check is carried out at the time a function is performed.** A supervisor observes the laboratory technician while he or she carries out the test.

There are two types of human redundancy that could be considered:

- **Active human redundancy** can be identified in human systems through some analogy with redundancy in hardware systems. Active human redundancy occurs when the individual performing a redundant function is involved in the task at hand; for instance, when two laboratory technicians take part in the analysis of one specimen.
- **Duplication, Overlap and Substitution.** Duplication exists when two different people perform the same function or if a reserve unit is present. Overlap exists when two people share some functional areas. For instance, when two people perform the same kind of laboratory testing, duplication takes place; however, when two people carry out different testing but they share the same equipment of parts of laboratory facilities refers to overlap. Substitution occurs when people rotate jobs.

Outcome based detection will take place once a test has been completed and the test results are checked for errors. As we have seen in the Manchester incident, detection of the radiologist's errors was done by colleagues of his who examined reports he had compiled. Double-checking is considered to be a standard practice in diagnostic services [Johnson and Patnick, 2000], although it was not mentioned in the other case studies.

There is little room for making recommendations here; this is primarily because outcome-based detection is most likely to occur outside the lab, since this is where test results are used—and thus evaluated. However, a number of options are applicable within a laboratory setting regarding detection through limiting functions—detection by constraints imposed on a diagnostic process. Some of these are discussed in the following section.

7.1.2 **User interface issues**

Human-Computer Interaction (HCI) plays a very important role in error handling due to the increasing reliance on software-automated testing. The user interface can influence both detection and further investigation, with an immediate impact on any problem/error indication communication that may take place.

Error detection

In the Sheffield and the Florida incident, user interface deficiencies resulted in a failure to detect software faults. Two conclusions can be drawn from these two incidents:

- **Critical calculations should be visible.** In Sheffield, the screening calculation had been removed from the user interface following users' request to reduce clutter. The result that was then indicated was a mere "high-risk" or "low-risk". Had the calculation been visible, and users would be able to perform outcome-based detection. This illustrates a trade-off that exists within usability. The key is that issues such as simplicity and readability had not adequately been considered in terms of the impact on safety – it is therefore necessary to take into account usability in hazard analysis and risk assessment.
- **Screening tests should be reported separately.** In Florida, the reporting system would only print out the positive patients, and negatives would be disregarded. This was done in a custom report which was also included other tests. The software fault resulted in the positives not being reported, but the printout report appeared normal as there were other test results present. The conclusion to be drawn from this is that critical tests should be reported separately, and there should be no filter applied. All test results should be included, but presented in an organized manner, with more information regarding the particular test that has been carried out.

The next section will consider how software applications could be used to improve screening error handling—primarily by facilitating further investigation.

7.1.3 **Application-level issues**

As we have seen in the case studies, problem detection occurred over time due to a discrepancy in the frequency of positives/negatives, based on the expectations of the nurse responsible (and primarily in the Sheffield and Florida incidents). Nurses were however unable to investigate themselves, while problem indication was ineffective—investigation activities to determine if there is an error were only possible within the lab.

It has been suggested that allowing for comparisons is a practical way of *evaluating outcomes* [Reason 1997; Rizzo et al., 1996]. Staff at the points of care where test results are used should be able to monitor and investigate trends of test results at their location. A potential approach to this is by automatically logging all test results reported from the lab, and calculating a mean of the entire set of test results through a software application. A graphical distribution of test values or results across time can help to notice potential discrepancies and support further investigation by helping to identify a timeframe within which the behaviour of test results is different than the one expected. The results of such a comparison, with a date and a number in the drop or rise of reported positives/negatives can stand as evidence for laboratory technicians to carry out further investigation.

Such an application would be particularly useful as it correlates historical data and represents it graphically. The nurses' expectations and any potential deviations would both be captured and documented.

The next section will discuss how communication during error handling could be improved. This relates to problem/error indication and further investigation.

7.2 Improving Communication

Improvement of communication during problem/error indication and during further investigation can be improved in various ways[15]. This section will present various interventions that have been suggested in related literature.

Non- technical interventions

- *Alter communication behaviours:* Such interventions focus on encouraging communication behaviours as a professional skill rather than as a personal style, and they are a matter of education and training.
- *Alter communication policies:* Mandatory policies should formalise certain aspects of communication, while there can be constraints on professional behaviour involving poor communication. This can be related to policies regarding incident reporting, and in the laboratory setting, reports made to the lab regarding possible errors detected in points of care.

Technical interventions

[15] Material from this section resulted in the following paper:

Chozos N, Wears RL and Perry S (2008). The role of communication in laboratory error handling. *Healthcare Ergonomics and Patient Safety (HEPS) conference*, Strasbourg, France 25-27 June.

With the merging of information and communication technologies, a number of different technical interventions have been suggested

- *Channels:* One of the simplest interventions is to introduce new communication channels, such as pagers, mobile phones, Internet, email and other new options for interaction. Such interventions may be very helpful, especially in teams which are geographically dispersed.

The following sections apply these recommendations to a diagnostic setting.

7.2.1 Problem/error indication

Non-technical interventions: Alter communication policies

A dedicated phone line

In many cases were the same nurse made several phone calls to a laboratory, different people picked up the phone. They were thus possibly not aware of previous reports, especially since log books for documenting abnormal test results were not used, or were not present at all. A dedicated phone line, and perhaps designated staff responsible for dealing with these phone calls, could greatly improve error handling for the following reasons:

- Formalization of incident reporting over the phone
- Presence of designated staff to deal with incident reporting
- Ability to log messages for further investigation purposes

This recommendation was also considered in a recent Scotland-wide study as a possibility to facilitate reporting by nurses [BBC news, 2009]. The study found that about

only about one third of reports resulted in some action taken, while the rest were largely ignored.

Double checking of critical values

Double checking of critical values can be seen as an intervention to support communication over the phone. This was recommended by a recent study which found that errors were significantly reduced [Barenfanger et al., 2004]. This requires that the person taking the result must read the result back to the lab as a check on correct communication and interpretation. Reduction of errors and better communication of important data can be achieved by asking all recipients (nurses, doctors, admin) to read back the message. This is a simple yet effective measure for laboratories to improve safety by minimizing the number of critical values missed. This however deals primarily with errors such as wrong patient name or other patient information, and miscommunication of a test result if it is done over the phone.

In this study, critical laboratory results were monitored. After receiving the message, the recipients of a telephoned message were asked to repeat the message. The recipients were asked to repeat the name of the patient, the test, and the result; the technologists noted this on the form. In addition, they noted the time necessary for the entire phone call and the extra time necessary to ask for the message to be repeated and for it to be repeated.

Out of a total 822 telephone contacts made for critical results, 29 errors were made (error rate, 3.5%). The major categories of errors were incorrect name of the patient, incorrect test result, incorrect specimen or test repeated, and refusal of the recipient to repeat the

message. The time required to deliver the message initially averaged 57.6 seconds per call. The time required to ask for the information to be repeated and for the recipient to repeat the message was a mean of 12.8 seconds per call. Times vary depending on the laboratory testing process. A call about a critical result from the microbiology laboratory inherently involves a more complex narrative than one would have in the chemistry laboratory.

There is an abundance of work that could be also considered here. For instance, Leonard et al. [2004] examine the role of communication in the effectiveness of teamwork, while Haig et al. [2006] consider a shared mental model for improving communication between clinicians. Such work could be integrated with the findings of this study in order to generate more detailed recommendations.

7.3 Risk calculation algorithmic issues

Error trapping is a common practice in software development, which involves detecting an error and producing an error message, taking some action on the erroneous result and either proceeding with execution or aborting the execution. This can either occur for run-time errors whose results are that are outside the defined range, or for infinity errors (e.g. division by zero).

The first type occurred in the Sheffield incident, while the second in the Florida incident (the other two incidents did not involve software). However, code error handling is not within the scope of this thesis, as it could be argued that this is still within error

prevention. This thesis focuses on the handling of laboratory errors, which assumes that software cannot always be reliable.

7.4 Job design and training

Training can be a key factor towards effective error handling [Chmiel and Wall, 1994]. There are implications for training for nurses and for laboratory staff. As a starting point, training should consider raising the awareness of problems that may lie in screening services [Chant et al., 2002]. Discussion of incidents and accidents in relation to medical processes and involved equipment can facilitate the understanding of cause-and-effect; the role of staff in the detection and recovery of such problems should also be party of that training.

A key aspect of training with regards to error handling would be to focus on the communication problems that have been found in this analysis. The expressions and terms used to transmit concerns over the phone to the laboratory technician can have a critical affect on a recovery process. This is an area that would require further research and is not considered in any depth here.

Overall, if any of the error handling recommendations suggested in this chapter were to be implemented, training should also cover their implementation. For instance, reporting over the phone should be part of medical staff's implicit training. This suggests that job design should primarily address error handling, with supportive training for a particular job specification.

Job design should encompass the appropriate portion of the responsibility for error recovery. One conclusion to be drawn from this analysis is what the capabilities of involved parties are in terms of detection and recovery. For instance, nurses that request, receive and use test results are limited to problem and error detection, with little ability to investigate the system. Therefore, their job specification should involve the monitoring of test results and the reporting of any concerns to their supervisor and/or the laboratory technician. In a similar way, the job specification of the laboratory technician should encompass following up reports and so on.

On the other hand, it is important not to over-formalize some of these processes. Creating additional tasks can impose a work overload, leading staff to find workarounds. Therefore, the level to which some of these recommendations should be introduced as part of everyday work and policy requires further research.

7.5 Chapter summary and discussion

This chapter outlined a number of high-level recommendations that may be applicable to healthcare screening services. They have partly been derived from the findings of accident analysis, while some additional recommendations are based on literature in error detection and HCI. These recommendations considered checking mechanisms, interface design, improvement of communication and training, while the relevance of each of these to the various stages of error recovery was also discussed.

The recommendations discussed here can be seen as preliminary, as they are not focused on a particular system and maintain a high level of abstraction; also, they lack validation. Further research would be required to instantiate and validate such recommendations.

Chapter 8: ER-STEP Validation and evaluation

A useful way for drawing conclusions regarding the applicability of an incident analysis approach is to distribute a scenario-based exercise to participants and compare their analysis and findings for consistency; such an activity was also done for ER-STEP. The exercise and the participant's findings can be found in Appendix B. The purpose of this section is to discuss some of the key issues that arose from the validation exercise.

8.1 Validation method overview

It should be mentioned that the aspect of ER-STEP that is subject to validation is the level to which analysts may consistently label events according to the error recovery framework stages. Otherwise, the set of activities that make up the method are identical to STEP, which has been widely accepted as a practical and straightforward technique.

In order to evaluate the level to which analysts may label error recovery events with consistency, an exercise along with a brief introduction to ER-STEP and an example of how it should be applied was given to four participants. The exercise is based on **Case Study 1:** Down's screening errors, Sheffield, so that the participants' results could be compared to the author's. This was done in two stages:

- **Stage 1:** *Initial evaluation by two experienced accident analysts*. At an early stage of the development of the method, it was necessary to get an expert opinion on the

feasibility of such an analysis, and to identify problems of technical nature. Therefore, the first draft of the technique, along with guidelines as to how it should be applied was given to two accident analysts and a scenario for them to analyze.

- **Stage 2:** *Final evaluation by two healthcare professionals*. Having revised the technique after its evaluation by the two experts, and having had further experience by analyzing the four incidents, the exercise was given to two healthcare professionals. It is important that people at the forefront are able to apply such a technique without necessarily having experience in accident analysis.

Stage 1: accident analysis experts

The initial analysis found there was not a sufficient distinction between communication events during further investigation and problem/error indication. During further investigation, a person will make an enquiry regarding a problem in order to take action themselves, or will instruct someone to carry out a specific activity. This is different to problem/error indication, where someone reports a problem/error, without having any control over the subsequent activities that are to take place. This was clarified in the section 3.5, where the error recovery framework is defined.

At the time of this evaluation, the technique was still under development and changed significantly since their exercise. The actual exercise and the produced findings are therefore not presented here.

Stage 2: Healthcare professionals

Table 8.1 illustrates the answers that the two healthcare professionals provided in the exercise.

Event	Participant 1	Participant 2
E1: Problem detection	√	√
E2: Problem indication	√	√
E3:Problem indication	√	√
E4:Further investigation	√	√
E5:Failure of further investigation	√	Failure of problem indication
E6: Further investigation	√	√
E7: Further investigation	√	√
E8:Problem indication	√	√
E9: Further investigation	√	√
E10: Problem indication	√	√
E11: Failure of further investigation	√	Failure of problem indication

Table 8.1: ER-STEP validation exercise findings.

The two participants only disagreed in two events; however, they were of the same type which appeared twice. Disagreement can therefore be placed only on one event (Or at least in terms of this methodology this was the only disagreement that was apparent. It is possible that they might have classified events the same way but for different reasons).

The disagreement was in the labelling of a 'failure of further investigation' event which was tagged as failure of 'problem indication' by the second participant. Following this, it was decided to brand all 'failure of' events in the same way: failure of error recovery, as

it may not always be clear where the failure was. As we can see in this case, the report of the nurse did not result in an effective investigation. It is therefore not straightforward to derive which of the two stages failed.

Comments from the two participants can also be found at the end of each of the two forms. Some of them were based on the lack of information regarding the incident, and in particular timing, which was a problem for the main analysis of the case studies anyway. This limitation cannot be placed on the technique, as it is only a matter of what information is available from data gathering.

Both participants commented on the possibility of further breaking down events, although the one acknowledged the fact that the analysis would become more "swamped". Further breakdown has occurred as a result of the analysis of the case studies. In addition, the purpose of the exercise was to consider whether the basic identification of events would be possible; further analysis should be up to the analyst.

An important comment was based on the fact that repetition of problem indication could be regarded as part of further investigation in order to confirm if that potential problem really exists. This is a rather challenging issue, as it is very difficult to understand the intentions the person who is initiating this communication. It could be assumed that this is subject to interpretation, especially as this analysis and the exercise are based on data gathering that others have performed. This could only be clarified if the interviews with involved actors targeted such issues.

8.2 Validation and evaluation findings

In order to claim that the technique can be widely applied with ease, further evaluation is required. However, it should be stated that this technique was developed for the purposes of the specific investigation into laboratory error handling; the development of the technique itself may have formed a research objective, however its complete validation is not an objective within the scope of this thesis. Nonetheless, the activities discussed in this section were very useful in making some considerable improvements and clarifications in the definitions of error recovery stages which may be subject to misinterpretation by the analyst.

Following the analysis of the exercises and the experience of applying the technique, some further conclusions can be drawn.

Benefits:

- The technique can be very useful in identifying, representing and communicating the activities that took place during error recovery during an incident.
- The reconstructed view can assist in identifying key problem areas of an error recovery process
- The focus on problem detection is particularly useful as it can help to reason about the role of "concerns" in error recovery. These can be the only error recovery initiating events in the absence of system feedback that can guide the investigator to identify errors.

Limitations:

- Still not always straightforward—but it has been found that different accident analysts may still conclude to different findings.
- Tool support—it has been very difficult to maintain the indexing and traceability between events in text, figures and STEP cards, while the drawing of figures had to be done with Microsoft Word, creating the possibility for inconsistencies in the use of colours, size of boxes etc.

8.3 Chapter summary

This chapter presented the approach that was undertaken towards validation of ER-STEP. The findings of the evaluation were also presented, while this section concluded with an evaluation of the technique as a result of the validation exercise and the experience accumulated with the application of ER-STEP with these four case studies.

Chapter 9: Evaluation and implications

This final chapter will present a summary and an evaluation of the research undertaken, before discussing the implications for practice and research this thesis has. The discussion about implications for practice will briefly state the relevance of the various findings that have emerged for different stakeholders, while directions for further research will concern patient safety, error detection and recovery, and accident analysis. This chapter will conclude with some final remarks.

9.1 Summary of research

The research presented here is an investigation into the factors that may impede detection and recovery of errors in screening tests. In order to identify and understand these factors and how they may inhibit an effective recovery, four incidents (three from the UK and one for the USA) were analysed with an analytical method that focuses on error handling activities that will take place. This method (ER-STEP) is an adaptation of STEP which has been integrated with a theoretical framework that illustrates the stages that make up an error recovery process from problem detection (initial concerns that something might be wrong) to error correction (chapters 3 and 4).

The findings of the four case studies (chapter 5) were integrated and compared to draw high-level conclusions about common problems in screening services in general. These conclusions (chapter 6) concerned both error recovery stages individually, and their relationship (e.g. how a specific kind of detection will result to a specific kind of further investigation). It was found that there are two key problem areas (which are anyway interrelated):

- **Problem detection is most likely to occur than error detection.** When there is a problem with a screening service, the first instance of detection is going to take place where test results are used. In almost all instances in the four case studies, this was done by the nurse responsible for requesting and received them. This detection is based on a discrepancy between the expected number of positives/negatives (probably per week or month).
- **Severe communication breakdowns throughout the recovery process:** Communication breakdowns will primarily occur because the reporting nurse will call the laboratory to report concerns, without convincing evidence that will motivate the laboratory technician to investigate further. However, these communication breakdowns are not only attributed to limited detection, but also to problematic—or a lack of—procedures for recording and handling complaints. The means of communication also play a role in the effectiveness of reporting of concerns. The informal reporting over the phone did not succeed in convincing laboratory technicians to investigate the system in order to find out if there is an error or not, whereas in the cases where face-to-face communication was possible, it was effective.

The sum of these findings resulted in the "laboratory error recovery" model, which classifies and correlates the various activities that may take place for the purpose of error handling outside and inside the laboratory.

These findings are used to generate recommendations for the improvement of detection and recovery (chapter 7). Additional literature was used from the areas of Human-Computer Interaction and systems engineering in order to provide some more detailed insight for technical and organizational interventions that aim at improving the preparedness of healthcare systems to detect and recover lab errors more efficiently. However these recommendations serve as a secondary aim of this thesis and would require more research in order to be further developed and validated.

Finally, Chapter 8 presents the validation and evaluation of ER-STEP. The exercise and the participants' responses can be found in Appendix B.

9.2 Evaluation of research

This section will discuss the contributions that this thesis has perhaps made to patient safety, accident analysis and error recovery, as well as the limitations and major problems that were faced during this work.

9.2.1 Contribution to practice

Screening programmes need to have adequate systems in place in order to "*to be able to respond to errors quickly*" [Screening Programme Director's report, 2005]. The several

screening failures that have been discussed throughout this thesis illustrate what can happen if screening programmes are not prepared to detect and recover from errors when they occur. While there are several studies in the area of laboratory error, they mainly focus on error types and frequencies. It has been argued that we have very limited understanding of the impact of laboratory errors and laboratory work in general on patient outcomes [Plebani and Carraro, 1997; Bonini et al., 2002]. These issues motivate this study. The findings and the approach employed can help to understand how healthcare systems could be better prepared to detect and deal with errors when they occur within screening services.

This study has identified a set of high-level conclusions—this level of abstraction was required in order to understand what the common key problem areas require that attention is placed in screening services in general (limited detection and communication barriers) and their impact on various stages of a recovery process. However, the application of these recommendations to a 'real world' environment would require a substantial amount of further technical analysis.

9.2.2 Contribution to accident and incident analysis

Although error handling is identified as a distinct area in the study of error, it is not considered as a separate issue in accident analysis. Up to this time, there is no technique available that focuses intrinsically on error detection and recovery. The adaptation of STEP to ER-STEP is a proposed approach to analyze the organizational response during a crisis. The validation and evaluation of the technique where already discussed in Chapter 8: in detail.

9.2.3 Cognitive science and systems engineering

The use of the error recovery framework as a means to analyze error handling activities has resulted in a better understanding of the relationship between such activities, and the variety of actions that may fall under a specific error recovery stage. For instance, we can better understand how different kinds of communication may be stimulated following the various kinds of detection, or what purposes further investigation has.

In addition, the inclusion of problem detection in the error recovery process has not—to this time—been done; problem detection, as a subject matter has been considered as an action of its own. In a similar manner, most recovery models and frameworks do not take into account the different mechanisms of detection (e.g. action-based, outcome-based, through limiting functions).

Finally, this study indicated an additional detection mechanism which had not been considered by Sellen [1994]: detection through investigation. This kind of detection suggests that not only the process of recovery, but also error detection may be organizational processes; especially in this context, where error detection through further investigation was inter-organizational.

9.2.4 Study limitations

Given the high-level view that is taken in this thesis, the recommendations require further input in order to be practically useful within a specific context. In such a case, it is possible some of the recommendations might not be applicable. In addition to this point, it is very difficult to validate the correctness and significance of the findings and

recommendations. This has perhaps been achieved to a certain extent by peer review publications, and especially the two in the Patient Safety and Ergonomics conference, as it has a clear focus on patient safety.

Another limitation is the hindsight bias [Johnson, 2003] with which the analyst views an accident. Although the focus on problem detection has been an important part of this thesis, it is unknown how many of reports based on concerns, without concrete evidence, are actually correct and not "noise". Nonetheless, this thesis has argued that events of "problem indication" should be at least recorded, as they can be of great importance during an incident investigation.

The final limitation here was the lack of data for the analysis of the Florida incident. The laboratory was unwilling to discuss the incident in detail, while the distance and the time-difference made follow-up discussions problematic from the UK. For these reasons, the analysis of the Florida incident were significantly more superficial in comparison to the UK incidents.

9.3 Implications for research

Research in error handling has been fairly limited, at least in comparison to error prevention. This section will discuss some potential directions for further research which consider error handling in healthcare, the application of ER-STEP in other domains, and the consideration of error handling within safety argumentation (safety cases).

9.3.1 Error handling in healthcare

Understanding the boundaries of plausibility

A combination of qualitative and quantitative methods should be employed to understand the thresholds (problem detection) that when passed would drive a nurse to report a potential problem (problem indication). This is based on the notion that false-yet-plausible test results are only plausible *in isolation*. The qualitative aspect would focus on examining a particular screening process, and, with a laboratory medicine-driven analysis, construct scenarios which describe potential instances of false-yet-plausible test results.

Focus groups consisting of nurses would then discuss these scenarios in order to explore the different levels of plausibility. Ideally, such a study should be run within a network of hospitals and GP practices that are all dependent on the same laboratory or radiology department, with the scenarios being based on that specific department. The findings of such a study could then be used to direct questionnaires that could be deployed on a larger scale, again focusing on how nurses would react to different levels of plausibility.

Hazard analysis

The focus of hazard analysis should aim to identify the causes of false test results, focusing on their potential plausibility. In order to do this, a multidisciplinary perspective on hazard analysis would be required, which would involve clinicians at points of care, nurses, laboratory technicians, quality and safety managers, and system and software engineers.

9.3.2 Application of ER-STEP in other domains

The error under investigation here can be seen to be information-based as it impacts medical test results; thus, safety implications arise after a considerable amount of time. It would be interesting to consider different kinds of errors and technical faults, as well as in other domains such as aviation and energy, where the timeframe within which recovery will have to occur is much shorter.

9.3.3 ER-STEP tool

Following the discussion regarding limitations of the technique, it would be very useful if there was tool support for ER-STEP. Tool support could also facilitate the restructuring of events to the error recovery focused view.

9.4 Final remarks

Error handling in screening programmes can be a complex, multi-departmental and inter-organizational process; while detection is most likely to occur where test results are used (and originally requested), rectification of the technical problem can only take place within the laboratory. In addition, the incident response can only be considered as complete once affected patients are all identified and contacted, as a misdiagnosis will continue to affect a patient as time passes. This part of error recovery may involve several other organizations, such as national and regulatory authorities, or independent auditors.

Problem detection was the initial recovery related activity in all of the incidents that were discussed here. Yet problem detection has neither been given enough attention in practice, nor in research. To an extent, this is understandable as problem detection often

occurs without the presence of an error, and the turbulence of medical environments along with limited time and resources do not allow for the investigation into any concern reported to the laboratory. Nonetheless, problem detection is critical as the plausibility of test results does not allow for error detection. The consideration of problem detection in the analysis of these four incidents has therefore been an important aspect of this thesis.

The technique proposed perhaps requires some further application in order to calibrate and better define the boundaries between activities that seem to overlap. In any case, this is the first proposition of an error recovery focused technique; the findings of the multiple analyses resulted in the screening error recovery model which could be used to design error recovery processes in a laboratory setting.

The involvement of physicians has been very important for the purposes of this work. A laboratory supervisor (Dr Frank Finley), a GP surgeon (Dr James Barnes) and an Emergency Medicine physician (Dr Robert Wears) have provided with very important insight, and with some evaluation and validation of the findings that resulted from this work.

To conclude, this thesis aimed at increasing our understanding of an important problem that medical practice currently faces. The primary contribution of this thesis is therefore seen to be the set of conclusions that were derived in Chapter 6, which can be used to design error recovery processes in NHS screening services. The author hopes that such work will be continued, and, in the long run, will contribute to the improvement of the

ability of the NHS to detect laboratory screening errors and better handle them, eventually offering safer and of higher quality services to individuals.

References

Advisory Committee on Breast Cancer Screening (2006). *Screening for Breast Cancer in England: Past and Future*, (NHSBSP Publication no 61) [Available at: http://www.cancerscreening.nhs.uk/breastscreen/publications/nhsbsp61.html, last accessed 2/6/2008].

Albers M (1999) Information Design Considerations for Improving Situation Awareness in Complex Problem-Solving. *Proceedings of the 17th annual international conference on Computer documentation*, New Orleans, Louisiana, United States pp: 154 – 158.

Allwood C (1984). Error detection processes in statistical problem solving. *Cognitive Science*, 8, pp. 413–437.

Amalberti R, Auroy Y, Berwick D, Barach P. Five system barriers to achieving ultra safe health care. Ann Intern Med. 2005;142:756-64.

Ash J, Berg M, Coiera E (2004). Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. *Journal of the American Medical Information Association*, 11, pp. 104-112.

Bagnara S and Rizzo A (1989). A Methodology for the Analysis of Error Processes in Human-Computer Interaction. *Proceedings of the Third International Conference on Human-Computer Interaction*, pp. 605-612.

Baker DP, Gustafson S, Beaubien J, Salas E, Barach P. Medical teamwork and patient safety: the evidence-based relation. Agency for Healthcare Research and Quality [retrieved 2008-02-28]. http:// www.ahrq.gov/qual/medteam

Bhasale AL, Miller GC, Reid SE and Britt HC (1998). Analysing potential harm in Australian general practice: an incident-monitoring study, *Medical Journal of Australia,* 169, pp. 73-76

Baker M (2003). Breast Cancer Services in Trafford and North Manchester: An Investigation into the Circumstances Surrounding a Serious Clinical Incident in Symptomatic Breast Services.
[Available at: www.northwest.nhs.uk/document_uploads/baker/baker_report.pdf, last accessed 20/4/2007].

Barenfanger J, Sautter RL, Lang DL, Collins MS, Hacek DM and Peterson, RL (2004). Improving Patient Safety by Repeating (Read-Back) Telephone Reports of Critical Information. *American Journal of Clinical Pathology*, 121, pp. 801-803.

Basnyat S, Chozos N, Johnson C, Palanque P (2005). Incident and Accident Investigation Techniques to Inform Model-Based Design of Safety-Critical Interactive Systems. *12th International workshop on Design, Specification and Verification of Interactive Systems*. *Newcastle upon Tyne, England.*

Bates D and Gawande A (2003). Improving Safety with Information Technology. *The New England Journal of Medicine*, 348, pp. 2526- 2534.

BBC news (2000). Down's Risk 'Misdiagnosed'.
[Available at: http://news.bbc.co.uk/1/hi/health/769640.stm, last accessed 22/4/2007]

BBC news (2001a). Down's Risk Error Report.
[Available at: http://news.bbc.co.uk/1/hi/health/1541661.stm, last accessed 22/4/2007]

BBC news (2001b), Millennium Bug blamed for test errors.

[Available at: http://news.bbc.co.uk/1/hi/health/1541557.stm, last accessed 2/6/2008]

BBC news (2002), Breast screening errors 'unacceptable'.
[Available at: http://news.bbc.co.uk/1/hi/health/1930485.stm, last accessed 2/6/2008]

BBC news (2003). Millennium Bug Blamed for Test Errors.
[Available at: http://news.bbc.co.uk/1/hi/health/1541557.stm, last accessed 22/4/2007]

BBC news (2005). Lab Error 'Led to shattered life'.
[Available at: http://news.bbc.co.uk/1/hi/england/4255205.stm, last accessed 22/4/2007]

BBC news (2009). Patient welfare fears 'ignored'.
[Available at: http://news.bbc.co.uk/1/hi/scotland/8042621.stm, last accessed 19/5/2009]

Benner L (1975). Accident Investigation: Multi-linear Events Sequencing Methods, *Journal of Safety Research*, 7 (2), pp 62-73.

Benner L (2003). Investigating Investigation Methodologies. *Workshop on Investigation and Reporting of Incidents and Accidents (IRIA)*, Williamsburg, Virginia, September 16-19.

Berg M (2004). *Health Information Management*. Routledge Health Management Series.

Blavier A, Rouy E, Nyssen A and De Keyser V (2005). Prospective Issues for Problem Detection. *Ergonomics*, pp. 758- 781.

Blumenthal D (1997). The Errors of Our Ways (Editorial). *Clinical Chemistry* 43 (8), p. 1305.

Bogner M (1994). Ed. *Human Error in Medicine*, Lawrence Erlbaum Associates. Hillsdale: New Jersey, USA.

Bove T and Andersen H (2003): Development of an Error Management Taxonomy in ATC. In: *Proceedings of the Tenth International Conference on Human-Computer Interaction*, pp. 407-411.

Bonini P, Plebani M, Ceriotti F and Rubboli F (2002). Errors in Laboratory Medicine. *Clinical Chemistry.*48, pp. 691-698

Brennan S. (2005). The NHS IT Project: *The Biggest Computer Program in the World... Ever!* Radcliffe Medical Press Ltd.

Burke M, Boal J, Mitchell R. A new look at the old. Communicating for better care: improving nurse-physician communication. Am J Nurs. 2004;104(12):40-7.

Burstein D. (2003) Canadian medical tests give reversed results. Risk Digest. [Available at: http://catless.ncl.ac.uk/Risks/23.19.html#subj2, last accessed 21/4/2007]

Buys JR and Clark LJ (1995). Events and Causal Factors Analysis (ECFA). Technical Research and Analysis Center, Idaho USA. [Available at: www.ceet.niu.edu/depts/tech/asse/tech482/trac14.pdf, last accessed 19/5/2009]

Chant S, Jenkinson T, Randle J, Russell G and Webb C (2002). Communication Skills Training in Healthcare: A Review of the Literature. *Nurse Education Today*, 22, pp. 189- 202.

Chmiel N and Wall T (1994). Fault Prevention, Job Design, and the Adaptive Control of Advanced Manufacturing Technology. *Applied Psychology: An International Review*, 42 (4), pp. 455- 473.

Chozos N (2008). Focusing accident analysis on error recovery: Three case studies in the NHS. *Journal of Reliability and Safety Engineering* [accepted for publication].

Chozos N, Wears RL and Perry S (2008). The role of communication in laboratory error handling. *Healthcare Ergonomics and Patient Safety (HEPS) conference*, Strasbourg, France 25-27 June.

Christian CK, Gustafson ML, Roth EM, Sheridan TB, Gandhi TK, Dwyer K, et al. A prospective study of patient safety in the operating room. Surgery. 2006;139:159-73.

Cicourel A (1990). The integration of distributed knowledge in collaborative medical diagnoses. in Galegher J, Krant, R and Egido C. Eds *Intellectual Teamwork: Social and Technological Foundations of Cooperative Work*, pp. 221-241.

Clarke DM (2005), Human Redundancy in Complex, Hazardous Systems: A theoretical framework. *Safety Science*. 43: pp. 566-677.

Commission for Health Improvement (2002). Investigation: West of London Breast Screening Service at Hammersmith Hospitals NHS Trust. [Available at: http://www.healthcarecommission.org.uk/_db/_documents/04005355.pdf, last accessed 17/4/2007]

Cowan DA (1986). Developing a Process Model of Problem Recognition. David A. Cowan. *The Academy of Management Review*, 11(4), pp. 763-776.

Dain S (2002). Normal Accidents: Human Error and Medical Equipment Design. *The Heart Surgery Forum*, 5(3), pp. 254- 257.

Davenport DL, Henderson WG, Mosca CL, Khuri SF, Mentzer RM, Jr. Risk-adjusted morbidity in teaching hospitals correlates with reported levels of communication and collaboration on surgical teams but not with scale measures of teamwork climate, safety climate, or working conditions. J Am Coll Surg. 2007;205:778-84.

De Boer A, Blommerde B, De Haas P, Sebek M, Lambregts-van Weezenbeek K, Dessens M and Van Soolingen D (2002). False-Positive Mycobacterium Tuberculosis Cultures in 44 Laboratories in the Netherlands (1993- 2000): Incidence, Risk Factors, and Consequences. *Journal of Clinical Microbiology*, 40(11), pp. 4004–4009.

Department of Health (2000). *An Organization with a Memory*. [Available at: http://www.dh.gov.uk/assetRoot/04/08/89/48/04088948.pdf] (Accessed at 28 Sept. 2006).

Department of Health (2001a). Health Department Announces Preliminary Results of St. Agnes Investigation. [Available at: http://www.dsf.health.state.pa.us/health/cwp/view.asp?A=190&Q=229407, last accessed 22/4/2007]

Department of Health (2001b). Health Department Orders St. Agnes Medical Center to Pay $447,500 Fine. [Available at: http://www.dsf.health.state.pa.us/health/cwp/view.asp?A=190&Q=229960, last accessed 22/4/2007]

Dix A, Finlay J, Abowd G and Beale R (1993). *Human-Computer Interaction-first edition*. Prentice Hall.

Dix A, Finlay J, Abowd G and Beale R (2004). *Human-Computer Interaction-third edition*. Pearson Education Limited. Harlow, England.

Davies G (1973). *Psychology of problem solving*. Basic Books, New York.

Dörner D and Schaub H (1994). Errors in Planning and Decision-making, and the Nature of Human Information Processing. *Applied Psychology: An International Review*, 42 (4), pp. 433- 453.

Dutke S (1994) Error Handling: Visualisations in the Human-Computer Interface and Exploratory Learning. *Applied Psychology: An International Review*, 42 (4), pp. 521- 541.

Edwards A and Elwyn G (2001). Understanding Risk and Lessons for Clinical Risk Communication about Treatment Preferences. *Quality and Safety in Health Care*, 10, pp. 9-13.

Farrell D and Peterson J (1982). Patterns of Political Behavior in Organizations. *The Academy of Management Review*, 7(3), pp. 403-412

Ferres et al. (2001), Report of the inquiry committee into the computer software error in. Downs syndrome screening, [Available at: www.sheffield.nhs.uk/resources/Downs-report.pdf, last accessed 20/6/2007]

Ferry T (1998). *Modern accident investigation and analysis*. Willey-IEEE, New York.

Frable W (2006). Surgical Pathology-Second Reviews, Institutional Reviews, Audits and Correlations. What's out There? Error or Diagnostic Variation*?* *Archive Pathology Laboratory Medicine*, 130, pp. 620- 624.

Frese, M (1991): Error Management or Error Prevention: Two Strategies to Deal with Errors in Software Design. *In: Proceedings of the Fourth International Conference on Human-Computer Interaction 1991*. pp. 776-782.

Frese M, Reason J and Zapf D (1994) Eds. Errors, Error Detection, and Error Recovery. Special Issue of *Applied Psychology: An International Review*. 42 (4).

Gaba D (2000), Structural and Organizational Issues in Patient Safety: A Comparison of Health Care to Other High-Hazard Industries. *Management Review*, 43: pp. 83-102.

Galegher J, Kraut RE, and Egido C (1990). Eds. *Intellectual teamwork: social and technological foundations of cooperative work.* Lawrence Erlbaum Associates, Inc. Mahwah, NJ, USA

Gaffney FA, Harden SW, Seddon R. Crew resource management: the flight plan for lasting change in patient safety. hcPro, Marbelhead; 2005.

Greenberg CC, Regenbogen SE, MD, Studdert DM, Lipsitz SR, Rogers SO, Zinner MJ, et al. Patterns of communication breakdowns resulting in injury to surgical patients. J Am Coll Surg. 2007;204:533-40.

Greenfield LJ. Doctors and nurses: a troubled partnership. Ann Surg. 1999;230:279-88.

Gillingham D, Blanco J and Lewko JH (1997). An Integrated Model for Error Management. *Disaster Prevention and Management*. 6 (3), pp. 186- 190.

Goldschmidt HMJ, Lent RW (1995). Gross errors and work flow analysis in the clinical laboratory. *Klin Biochem Metab*, 3, pp.131–140.

Greenwell SW, Strunk E and Knight JC (2004). *Failure Analysis and the Safety-Case Lifecycle*. Human Error, Safety and Systems Development, pp.163-176.

Hendrick K, Benner L (1987). *Investigating Accidents with Sequentially Timed and Events Plotting (STEP)*: Marcel Decker: New York.

Hickney CM (1997). Defining the Best Quality Control Systems for Design and Inspection. *Clinical Chemistry,* 43, pp. 873-879.

Howanitz T (2005). Errors in Laboratory Medicine: Practical Lessons to Improve Patient Safety. *Archives of Pathology and Laboratory Medicine,* 129, pp. 1252- 1261.

Haig KM, Sutton S, Whittington J. SBAR: a shared mental model for improving communication between clinicians. Jt Comm J Qual Patient Saf. 2006; 32 (3):167-7

Institute of Medicine (1999*). To err is human: building a safer health system*. Washington, DC: National Academies Press.

Jambon F (1997). Error Recovery Representations in Interactive System Development. *Third Annual ERCIM Workshop on "User Interfaces for All",* Obernai, France, 3-4 November, pp. 177- 182.

Jambon F (1998). Taxonomy for Human Error and System Fault Recovery from the Engineering Perspective. *International Conference on Human- Computer Interaction in Aeronautics (HCI- Aero'98)*, Montreal, Canada, 27- 29 May, pp. 55- 60.

Jenny RW and Tarentino KY (2000). Causes of Unsatisfactory Performance in Proficiency Testing. *Clinical Chemistry,* 46 (1), pp. 89-99.

Johnson, WG (1973). *MORT - The Management Oversight and Risk Tree.* SAN 821-2. US Atomic Energy Commission.

Johnson CW (2002). The Causes of Human Error in Medicine. *Cognition, Technology and Work*. 4, pp. 65-70.

Johnson CW (2003). *Failure in Safety-Critical Systems: A Handbook of Accident and Incident Reporting*, University of Glasgow Press, Glasgow, Scotland,

Johnson CW (2004). Communication Breakdowns Between the Suppliers and Users of Clinical Devices. *Biomedical Instrumentation and Technology*. 38(1): p. 54-78.

Johnson J and Patnick J (2000). Achievable standards, Benchmarks for reporting, and Criteria for evaluating cervical cytopathology. *NHSCSP Publication No 1*

Keil M (2001). Blowing the Whistle on Troubled Software Projects. *Communications of the ACM*. 44, 4, pp. 87-90.

Klein G, Pliske R, Crandall B and Woods D (2005). Problem Detection. *Cognition, Technology and Work*, 7, pp. 14-28.

Kontogiannis T (1997). A Framework for the Analysis of Cognitive Reliability in Complex Systems: A Recovery Centred Approach. *Reliability Engineering and System Safety*, 58, pp. 233-248.

Lapworth R and Teal TK (1994). Laboratory blunders revisited. Annals of Clinical Biochemistry, 31, pp.78-84.

Leveson N (2001). NASA technical report. Evaluating Accident Models using Recent Aerospace Accidents (Part 1: Event-Based Models) [Available at: http://sunnyday.mit.edu/accidents/nasa-report.pdf, last accessed 2/6/2008].

Leveson N, Daouk M, Dulac N and Marais K (2003). Applying STAMP in Accident Analysis, *Workshop on Investigation and Reporting of Incidents and Accidents (IRIA)*, Virginia, September 16-19.

Lewis C and Norman DA (1986). Designing for Error, in Norman DA and Draper SW. *Eds, User Centred Systems Design, Lawrence Erlbaum, Hillsdale*, NJ, pp. 411-432.

Liang B (2002). A System of Medical Error Disclosure. *Quality and Safety in Health Care*, 11, pp. 64-48.

Leonard M, Graham S, Bonacum D. The human factor: The critical importance of effective teamwork and communication in providing safe care. Qual Saf Health Care. 2004;13 (Suppl 1):i85-i90.

Lingard L, Espin S, Whyte S, Regehr G, Baker GR, Reznick R, et al. Communication failures in the operating room: an observational classification of recurrent types and effects. Qual Saf Health Care. 2004;13:330-4.

McFerran S, Nunes J, Pucci D, Zuniga A. Perinatal patient safety project: a multicenter approach to improve performance reliability at Kaiser Permanente. J Perinat Neonatal Nurs. 2005;19:37-45.

Mackie J and Sommerville I (2000). Failures of healthcare systems. *First Dependability IRC Workshop.* Edinburgh, United Kingdom.

Near P and Miceli M (1985). Organizational Dissidence: The Case of Whistle-Blowing. *Journal of Business Ethics*, 4, pp. 1-16.

Newman WM and Lamming MG (1995). *Interactive System Design.* Addison-Wesley, Oxford, UK.

Norman D (1983). Design Rules Based on Analyses of Human Error. *Communications of the ACM*, 26 (4), pp. 254-258.

Norman D (1990). The Problem of Automation: Inappropriate Feedback and Interaction, Not Over- Automation. *Philosophical Transactions of the Royal Society of London*, B, pp. 1-18.

Nutting PA, Main DS, Fischer PM, Stull TM, Pontious M, Seifert M, et al.. (1996) Problems in laboratory testing in primary care. *Journal of the American Medical Association,*275, pp.635–639.

NUREG-0600 (1980). Investigation into the March 28, 1979 Three Mile Island Accident by the Office of Inspection and Enforcement.

Pagliari C, Gilmour M and Sullivan F (2004). Electronic Clinical Communications Implementation (ECCI) in Scotland: A mixed-methods Programme Evaluation. *Journal of Evaluation in Clinical Practice*. 10 (1), pp. 11-20.

Pai S (2005). Laboratory Tests: Proper Communication Reduces Error. *Student British Medical Journal*, 13, pp. 397-440.

Plebani M and Carraro P (1997). Mistakes in a stat Laboratory: Types and Frequency. *Clinical Chemistry,* 43(8), pp. 1348-1351.

Programme Director's report (2005). Screening in England Summer 2003 – Spring 2004 [Available at: http://www.library.nhs.uk/screening/ViewResource.aspx?resID=64999, last accessed 2/6/2008].

Patterson ES, Roth EM, Woods DD, Chow R, Gomes JO, et al. Handoff strategies in settings with high consequences for failure: lessons for health care operations. Int J Qual Health Care. 2004;16:125-32.

Rasmussen J (2003). The Role of Error in Organizing Behaviour. *Quality and Safety in Health Care*, 12, pp. 377-383.

Reason J (1990). *Human Error*. New York: Cambridge University Press.

Reason J (1997). *Managing the Risks of Organizational Accidents*. Aldershot: Ashgate.

Reason J (2000). Human Error: Models and Management. *British Medical Journal*, 320, pp. 768- 770.

Reason J (2004). Beyond the Organizational Accident: The Need for "Error Wisdom" on the frontline. *British Medical Journal: Quality and Safety in Healthcare*, 13, pp. 28-33.

Rizzo A, Parlangeli O, Marchigiani E and Bagnara S (1996). The Management of Human Errors in User- Centred Design. *SIGCHI Bulletin*, 28 (3), pp. 114- 118.

Schiff G (2006). Ed. *Getting Results: Reliably Communicating and Acting on Critical Test Results*. Joint Commission Resources, Illinois.

Schoop M (1999). An Empirical Study of Multidisciplinary Communication in Healthcare using a Language-Action Perspective. *Fourth International Workshop on the Language Action Perspective on Communication Modeling (LAP 99)*, Jonkoping, pp. 59-72.

Sellen AJ (1994). Detection of Everyday Errors. *Applied Psychology: An International Review*, 42 (4), pp. 475-298.

Shepherd M (2004). Improving Healthcare Systems Following an Incident Investigation. *Proceedings of the 26th Annual International Conference of the IEEE EMBS, San Francisco, California, USA, September 1-5,* pp. 3500- 3502.

Shuren A (2000). Health care delivery errors: Patient safety falls prey to politics. Journal of Pediatric Health Care, 15 (1), pp. 42 - 44.

Sirota LS (2005). Error and Error Reduction in Pathology. *Archives of Pathology and Laboratory Medicine,* 129, pp. 1228-1233.

Skantze G (2003). Exploring Human Error Handling Strategies, ISCA Workshop on Error Handling in Spoken Dialogue Systems, August 28-31, 2003 Château d'Oex, Vaud, Switzerland.

Smith, BJ and McNeely DD (1999). The Influence of an Expert System for Test Ordering and Interpretation on Laboratory Investigations. *Clinical Chemistry*, 45 (8), pp. 1168-1175.

Solet DJ, Norvell JM, Rutan GH, Frankel RM. Lost in translation: Challenges and opportunities in physician-to-physician communication during patient handoffs. Academic Medicine. 2005;80(12): 1094-9.

Stahl M, Lund ED, Brandslund I (1998). Reasons for a laboratory's inability to report results for requested analytical tests. *Clinical Chemistry*, 44, 2195–2197.

Standards Unit, Evaluation and Standards Laboratory (2008). Quality Assurance in the diagnostic Serology and Virology Laboratory. [Available at: http://www.hpa-

standardmethods.org.uk/documents/qsop/pdf/qsop27.pdf, last accessed 07-Oct-08].

Sutcliffe K, Lewton E, Rosenthal M. Communication failures: an insidious contributor to medical mishaps. Academic Medicine. 2004;79: 186-94.

The Register (2001). Down's Syndrome screening failures linked to Y2K bug http://www.theregister.co.uk/2001/09/14/downs_syndrome_screening_failures_linked/

Thomas EJ, Sexton JB, Helmreich RL. Discrepant attitudes about teamwork among critical care nurses and physicians. Crit Care Med. 2003;31:956-9.

Vaughan D (1996). *The Challenger Launch Decision. Risky Technology, Culture, and Deviance at NASA*. The University of Chicago University Press.

Viller S, Bowers J and Rodden T (1999). Human Factors in Requirements Engineering: A Survey of Human Sciences Literature relevant to the Improvement of Dependable Systems Development Processes. *Interacting with Computers*, 11, pp. 665- 698.

Watt Committee (1988), *Chernobyl Accident and its Implications for the United*. Spon Press, London UK.

Wears R (2004). More on Computer Glitches and Laboratory Result Reporting. *The Risk Digest*, 24 (64), [Available at: http://catless.ncl.ac.uk/Risks/23.64.html%23subj4#subj4, last accesses at: 19/5/09]

Wears R (2005) (Editorial). Computer Technology and Clinical Work. *Journal of the American Medical Association*, 2005. 293(10): p. 1261-1233.

Wears R, Perry S, Chozos N and Johnson C (2006). Plausible but Wrong: A Failure Phenotype in Health IT. [presentation]. *AHRQ 2006 Annual Patient Safety and Health IT Conference, Washington, June 4-7.* [Available at: www.ahrq.gov. last accessed at: 10/4/2007]

Weick KE, Sutcliffe KM. Managing the unexpected: assuring high performance in an age of complexity. San Francisco: Jossey-Bass; 2001

West E. Organizational sources of safety and danger: sociological contributions to the study of adverse events. Quality in Health Care. 2000;9:120-6.

Wickens and McCarley (2007). *Applied Attention Theory*. CRC Press.

Wilmot S (2000). Nurses and Whistle-blowing: The Ethical Issues. *Journal of Advanced* Nursing, 32(5), pp. 1051-1057.

Winograd, F and Flores H (1986). *Understanding computers and cognition*. Ablex Publishing Corp. Norwood, USA.

Witte DL, VanNess SA, Angstadt DS and Pennell BJ (1997). Errors, mistakes, blunders, outliers, or unacceptable results: how many? *Clinical Chemistry,* 43:1352-1356.

Woods D (1987). Commentary: Cognitive Engineering in Complex and Dynamic Worlds. *International Journal of Man-Machine Studies* 27(5-6): 571-585

Yang Y (1992). Anatomy of the design of an undo support facility. *International Journal of Man-Machine Studies*, 1, pp. 81-95.

Yuan S, Astion ML, Schapiro J and Limaye AP (2005). Clinical Impact Associated with Corrected Results in Clinical Microbiology Testing. *Journal of Clinical Microbiology*, pp. 2188-2193.

Zapf D, Maier GW, Rappensperger G and Irmer C (1994). Error Detection, Task Characteristics, and Some Consequences for Software Design. *Applied Psychology: An International Review*, 42 (4), pp. 499- 520.

Appendix A: STEP cards

A.1 Incident 1: Down's screening errors, Sheffield

The data source for this incident is the formal Sheffield Inquiry report (Ferres et al., 2001). The 'data source' cell in each of the following STEP cards will refer to the specific paragraph number (description of events can be found in section 10, pp. 63-74 of the Report). In addition, the actors' names have been disclosed; in the Report they have been given aliases which will also be used here. Finally, there are several occasions where some information (in most cases the time/data event began and duration) has not been explicitly mentioned in the Report; therefore assumptions had to be made. Where necessary, these are highlighted by *italic* fonts in the STEP cards.

Event card id: **Event 1.1**	
Actor:	Dr T
Action:	**No classification**
Event location:	Immunology Department
Time/date event began:	7th December, 1999
Event duration:	
Data source:	10.2
Description:	Issues a draft of a document describing a new incident reporting scheme

Event card id: **Event 1.2**	
Actor:	*Dr B*
Action:	**No classification**
Event location:	Immunology Department
Time/date event began:	*1st January, 2000*
Event duration:	
Data source:	10.2
Description:	Introduces the new incident reporting scheme

Event card id: **Event 1.3**	
Actor:	Mr R
Action:	No classification
Event location:	Immunology Department
Time/date event began:	4th January
Event duration:	
Data source:	10.3
Description:	Undertakes Y2K testing on PathLan.

Event card id: **Event 1.4**	
Actor:	MGL sister
Action:	**Problem detection**
Event location:	Hospital B
Time/date event began:	*Mid January*
Event duration:	*Approx. 2 weeks*
Data source:	10.3
Description:	Becomes concerned as the number of positives she had received was lower than expected

Event card id: **Event 1.5**	
Actor:	MGL sister
Action:	**Problem indication**
Event location:	Hospital B
Time/date event began:	End of January
Event duration:	
Data source:	10.3
Description:	Reports concerns to immun. Dept. over the phone.

Event card id: **Event 1.6**	
Actor:	MGL sister
Action:	**Error detection**
Event location:	Hospital B
Time/date event began:	Mid January
Event duration:	
Data source:	10.3
Description:	Notices the results for an older woman are “unrealistically low”

Event card id: **Event 1.7**	
Actor:	MGL sister
Action:	**Problem indication**
Event location:	Hospital B
Time/date event began:	End of January
Event duration:	
Data source:	10.3
Description:	Reports concerns [Event 1.4] to immunology department over the phone.

Event card id: **Event 1.8**	
Actor:	MGL sister
Action:	**Error indication**
Event location:	Hospital B
Time/date event began:	End of January
Event duration:	
Data source:	10.3
Description:	Reports [Event 1.6] to Immun. Dept. over the phone.

Event card id: **Event 1.9**	
Actor:	Ms S
Action:	**Further investigation**
Event location:	Immunology Department
Time/date event began:	End of January
Event duration:	
Data source:	10.3
Description:	Asks Mr R "colleague who had undertaken Y2K testing [Event 1.3] about event 1.8.

Event card id: **Event 1.10**	
Actor:	Mr R
Action:	**Failure of Further investigation**
Event location:	Immun. Dept.
Time/date event began:	End of January
Event duration:	
Data source:	10.3
Description:	Reassures Ms S that there is no Y2K problem with PathLan.

Event card id: **Event 1.11**	
Actor:	Ms S
Action:	**Failure of Further investigation**
Event location:	Immun. Dept.
Time/date event began:	End of January
Event duration:	
Data source:	10.3
Description:	Does not log incident [Events 1.7 and 1.8]

Event card id: **Event 1.12**	
Actor:	MGL sister
Action:	**Further investigation**
Event location:	Hospital B
Time/date event began:	End of January
Event duration:	*2 months*
Data source:	10.4
Description:	Monitors her own screen positive results

Event card id: **Event 1.13**	
Actor:	MGL sister
Action:	**Problem indication**
Event location:	Hospital B
Time/date event began:	*April*
Event duration:	
Data source:	10.5
Description:	Reports concerns over the phone

Event card id: **Event 1.14**	
Actor:	<unknown> someone in immunology department
Action:	**Failure of Problem indication**
Event location:	Immunology department
Time/date event began:	April
Event duration:	
Data source:	10.5
Description:	Reassures that there is no problem with Down' screening.

Event card id: **Event 1.15**	
Actor:	Mr K
Action:	**No classification**
Event location:	Immunology department
Time/date event began:	April 2000
Event duration:	
Data source:	10.6
Description:	Requests from Dr A for a routine audit to be carried out for CPA visit due April 18th.

Event card id: **Event 1.16**	
Actor:	Dr A
Action:	**No classification**
Event location:	Immunology department
Time/date event began:	April 2000
Event duration:	
Data source:	10.6
Description:	In response to [Event 1.16], Dr A suggests

	the data would not be available but he would be able to respond to any queries made during CPA visit,

Event card id: **Event 1.17**	
Actor:	Mr K
Action:	**No classification**
Event location:	Immunology department
Time/date event began:	17th April
Event duration:	
Data source:	10.7
Description:	Performs random check of the 'High book'.

Event card id: **Event 1.18**	
Actor:	Mr K
Action:	**Problem detection**
Event location:	Immunology department
Time/date event began:	17th April
Event duration:	
Data source:	10.9
Description:	Notices there are too many positives reported in High book.

Event card id: **Event 1.19**	
Actor:	Mr K
Action:	**Further investigation**
Event location:	Immunology department
Time/date event began:	April 17th
Event duration:	
Data source:	10.10
Description:	Instructs Mr M to audit the screen positive rate

Event card id: **Event 1.20**	
Actor:	Mr M
Action:	**No classification**
Event location:	Immunology department
Time/date event began:	April 17th
Event duration:	
Data source:	10.10
Description:	Assumes 'human error'

	has been made during maintenance of the High book.

Event card id: **Event 1.21**	
Actor:	Mr M
Action:	**Failure of further investigation**
Event location:	Immunology department
Time/date event began:	April 17th
Event duration:	
Data source:	10.10
Description:	Does not consider this [Event 1.19] as urgent

Event card id: **Event 1.22**	
Actor:	Mr K
Action:	**No classification**
Event location:	Immunology department
Time/date event began:	April 19th
Event duration:	18 days
Data source:	10.11
Description:	Goes on leave until 8th of May

Event card id: **Event 1.23**	
Actor:	Midwife coordinator
Action:	**Problem detection**
Event location:	Hospital C
Time/date event began:	13th April
Event duration:	*2 weeks*
Data source:	10.13
Description:	Becomes concerned as no high-risk test results have been received

Event card id: **Event 1.24**	
Actor:	Midwife coordinator
Action:	**Problem indication**
Event location:	Hospital C
Time/date event began:	2nd May
Event duration:	
Data source:	10.13
Description:	Reports concerns

	[Event 1.23] to Mr M over the phone.

Event card id: **Event 1.25**	
Actor:	Mr M
Action:	**Failure of problem indication**
Event location:	Immunology department
Time/date event began:	2nd May
Event duration:	
Data source:	10.13
Description:	Does not consider Event 1.24 serious enough to notify Dr B.

Event card id: **Event 1.26**	
Actor:	Midwife coordinator
Action:	**Problem indication**
Event location:	Hospital C
Time/date event began:	*Week commencing 2nd May*
Event duration:	
Data source:	10.14
Description:	Reports concerns again [Event 1.23], this time to Ms S (over the phone).

Event card id: **Event 1.27**	
Actor:	Ms S
Action:	**No classification**
Event location:	Immunology department
Time/date event began:	Same day as Event 1.26
Event duration:	
Data source:	10.14
Description:	Replies she will notify Mr M regarding [Event 1.26]

Event card id: **Event 1.28**	
Actor:	Ms S
Action:	**Failure of Problem indication**
Event location:	Immunology department
Time/date event began:	Same day as Event 1.26
Event duration:	Until 17th May
Data source:	10.15
Description:	*Does not inform Mr M*

Event card id: **Event 1.29**	
Actor:	Ms S
Action:	**Failure of Problem indication**
Event location:	Immunology department
Time/date event began:	Same day as Event 1.26
Event duration:	Until 17th May
Data source:	10.15
Description:	Does not respond to [Event 1.26]

Event card id: **Event 1.30**	
Actor:	Midwife coordinator
Action:	**Problem indication**
Event location:	Hospital C
Time/date event began:	17th May
Event duration:	
Data source:	10.15
Description:	Reports to Dr B that she hasn't received a high-risk pregnancy report for 5 weeks, although she would expect 5-10 per week.

Event card id: **Event 1.31**	
Actor:	Dr B
Action:	**No classification**
Event location:	Immunology department
Time/date event began:	17th May
Event duration:	
Data source:	10.15
Description:	Responds she will get back to her (midwife coordinator) regarding Event 1.30.

Event card id: **Event 1.32**	
Actor:	Midwife coordinator
Action:	**Further investigation**
Event location:	Hospital C

Time/date event began:	17th May
Event duration:	
Data source:	10.15
Description:	Consults with the Head of Midwifery regarding what actions to take in order to pursue a resolution of her concerns.

Event card id: **Event 1.33**	
Actor:	Midwife coordinator
Action:	**Further investigation**
Event location:	Hospital C
Time/date event began:	17th May
Event duration:	
Data source:	10.16
Description:	Writes a letter to the immunology department addressed to Mr K

Event card id: **Event 1.34**	
Actor:	Midwife coordinator
Action:	**Problem indication**
Event location:	Hospital C
Time/date event began:	17th May
Event duration:	
Data source:	10.16
Description:	Writes a letter to the immunology department addressed to Mr K

Event card id: **Event 1.35**	
Actor:	Midwife coordinator
Action:	**Further investigation**
Event location:	Hospital C
Time/date event began:	17th May
Event duration:	
Data source:	10.17
Description:	Writes a letter to the immunology department addressed to Mr K

Event card id: **Event 1.36**	
Actor:	Dr B
Action:	**Further investigation**
Event location:	Immunology department
Time/date event began:	17th May
Event duration:	
Data source:	10.17
Description:	Has a discussion with Ms P regarding Event 1.34.

Event card id: **Event 1.37**	
Actor:	Ms P
Action:	**Further investigation**
Event location:	Immunology department
Time/date event began:	17th May
Event duration:	
Data source:	10.17
Description:	Informs Dr B Mr K had asked Mr M to undertake an audit earlier that month (May).

Event card id: **Event 1.38**	
Actor:	Mr M
Action:	**No classification**
Event location:	Immunology department
Time/date event began:	*Beginning of May*
Event duration:	Until 17th May
Data source:	10.17
Description:	Mr M does not perform audit.

Event card id: **Event 1.39**	
Actor:	Dr B
Action:	**Further investigation**
Event location:	Immunology department
Time/date event began:	17th May
Event duration:	
Data source:	10.17
Description:	Requests audit to be performed immediately,

	and have the results on his desk.

Event card id: **Event 1.40**	
Actor:	*Mr M*
Action:	**Further investigation**
Event location:	Immunology department
Time/date event began:	17th May
Event duration:	2 days
Data source:	10.17
Description:	Performs audit

Event card id: **Event 1.41**	
Actor:	Antenatal staff
Action:	**Error detection** *(2) [(2) is required as this is a-what seemed to be-different error]*
Event location:	Antenatal care
Time/date event began:	17th May
Event duration:	
Data source:	10.20
Description:	Notice DoB in two Down's screening reports are wrong.

Event card id: **Event 1.42**	
Actor:	Antenatal staff
Action:	**Error indication (2)**
Event location:	Antenatal care
Time/date event began:	17th May
Event duration:	
Data source:	10.20
Description:	Report event 1.41 to Ms S, immunology department.

Event card id: **Event 1.43**	
Actor:	Ms S
Action:	**No classification (2)**
Event location:	Immunology department
Time/date event began:	17th May
Event duration:	
Data source:	10.21
Description:	Attempts to change DoBs for the two reports in event 1.42.

Event card id: **Event 1.44**	
Actor:	Ms S
Action:	**Error detection**
Event location:	Immunology department
Time/date event began:	17th May
Event duration:	
Data source:	10.21
Description:	Notices risk calculation remains unaltered following event 1.43.

Event card id: **Event 1.45**	
Actor:	MS S
Action:	**Error indication**
Event location:	Immunology department
Time/date event began:	17th May
Event duration:	
Data source:	10.23
Description:	Reports event 1.44 to Mr L.

Event card id: **Event 1.46**	
Actor:	Mr L
Action:	**Error indication**
Event location:	Immunology department
Time/date event began:	17th May
Event duration:	
Data source:	10.23
Description:	Phones Mr W at Hartlepool.

Event card id: **Event 1.47**	
Actor:	Mr W
Action:	**No classification**
Event location:	Hartlepool hospital
Time/date event began:	17th May
Event duration:	
Data source:	10.23
Description:	Mr W is absent

Event card id: **Event 1.48**	
Actor:	Mr L
Action:	**Error indication**
Event location:	Immunology department
Time/date event began:	17th May
Event duration:	
Data source:	10.23
Description:	Leaves a voice-mail for Mr W.

Event card id: **Event 1.49**	
Actor:	Mr M
Action:	**No classification**
Event location:	Immunology department
Time/date event began:	18th May
Event duration:	
Data source:	10.20
Description:	Places a note on Dr B's desk writing "1.7%" following event 1.40.

Event card id: **Event 1.50**	
Actor:	Dr B
Action:	**No classification**
Event location:	Immunology department
Time/date event began:	7.00pm ,18th May
Event duration:	
Data source:	10.17
Description:	Sees note (event 1.41) on his desk

Event card id: **Event 1.51**	
Actor:	Dr B
Action:	**Failure of further investigation**
Event location:	Immunology department
Time/date event began:	7.00pm ,18th May
Event duration:	
Data source:	10.17
Description:	Does not pay attention to that note *(based on assumption that he did not understand what that note meant).*

Event card id: **Event 1.52**	
Actor:	Mr K
Action:	**No classification**
Event location:	Immunology department
Time/date event began:	Friday, 19th May
Event duration:	
Data source:	10.18
Description:	Returns to work

Event card id: **Event 1.53**	
Actor:	Dr B
Action:	**No classification**
Event location:	Immunology department
Time/date event began:	Friday, 19th May
Event duration:	
Data source:	10.18
Description:	Is absent as his wife goes into labour.

Event card id: **Event 1.54**	
Actor:	Dr B
Action:	**Problem indication**
Event location:	Immunology department
Time/date event began:	19th May
Event duration:	
Data source:	10.18
Description:	Speaks on the phone with Mr K to confirm that the results are locked in his office.

Event card id: **Event 1.55**	
Actor:	Mr K
Action:	**Further investigation**
Event location:	Immunology department
Time/date event began:	19th May
Event duration:	
Data source:	10.18
Description:	Agrees to look into the matter.

Event card id: **Event 1.56**	
Actor:	Midwife coordinator
Action:	**Further investigation**
Event location:	Immunology department
Time/date event began:	Evening, 19th May
Event duration:	
Data source:	10.19
Description:	Phones Mr K regarding the letter she sent (event 1.33)

Event card id: **Event 1.57**	
Actor:	Mr K
Action:	**Further investigation**
Event location:	Immunology department
Time/date event began:	22nd May
Event duration:	
Data source:	10.26
Description:	Receives letter through fax (event 1.33).

Event card id: **Event 1.58**	
Actor:	Dr B
Action:	**No classification**
Event location:	Immunology department
Time/date event began:	22nd May
Event duration:	1 day
Data source:	10.26
Description:	Away on paternity leave

Event card id: **Event 1.59**	
Actor:	Dr B
Action:	**No classification**
Event location:	Immunology department
Time/date event began:	11:30am, 23rd May
Event duration:	7 hours
Data source:	10.27
Description:	Is present at hospital (but not in immunology department).

Event card id: **Event 1.60**	
Actor:	Mr K
Action:	**Further investigation**
Event location:	Immunology department
Time/date event began:	23rd May
Event duration:	
Data source:	10.28
Description:	Asks Mr M to provide him with the results of the audit.

Event card id: **Event 1.61**	
Actor:	Mr K
Action:	**Further investigation**
Event location:	Immunology department
Time/date event began:	23rd May
Event duration:	
Data source:	10.28
Description:	Examines audit results

Event card id: **Event 1.62**	
Actor:	Mr K
Action:	**Error detection**
Event location:	Immunology department
Time/date event began:	23rd May
Event duration:	
Data source:	10.28
Description:	Notices only 2% positives had been reported since January.

Event card id: **Event 1.63**	
Actor:	Mr K
Action:	**Further investigation**
Event location:	Immunology department
Time/date event began:	23rd May
Event duration:	"some time"
Data source:	10.28
Description:	Examines analytical values

Event card id: **Event 1.64**	
Actor:	Mr K
Action:	**Further investigation**
Event location:	Immunology department
Time/date event began:	23rd May
Event duration:	
Data source:	10.28
Description:	Urges Mr L to contact Mr W.

Event card id: **Event 1.65**	
Actor:	Mr L
Action:	**Error indication**
Event location:	Immunology department
Time/date event began:	Morning, 24th May
Event duration:	
Data source:	10.28
Description:	Contacts Mr W, Hartepool over the phone.

Event card id: **Event 1.66**	
Actor:	Mr W
Action:	**Further investigation**
Event location:	Hartlepool
Time/date event began:	9:30am, 24th May
Event duration:	Approx. 30'
Data source:	10.29
Description:	Examines PathLan

Event card id: **Event 1.67**	
Actor:	Mr W
Action:	**Error explanation**
Event location:	Hartlepool
Time/date event began:	24th May
Event duration:	*A few minutes*
Data source:	10.29
Description:	Identifies bug

Event card id: **Event 1.68**	
Actor:	Mr W
Action:	**Error correction**
Event location:	Hartlepool
Time/date event began:	24th May
Event duration:	*A few minutes*
Data source:	10.29
Description:	Removes bug

Event card id: **Event 1.69**	
Actor:	Mr K
Action:	**Further investigation**
Event location:	Immunology department
Time/date event began:	24th May
Event duration:	
Data source:	10.30
Description:	Informs Dr A that he will find all high-risk cases that have been wrongly reported (face-to-face)

Event card id: **Event 1.70**	
Actor:	Mr K
Action:	**Further investigation**
Event location:	Immunology department
Time/date event began:	Morning, 24th May
Event duration:	*Approx. 12 hours (until 11.30pm).*
Data source:	10.34
Description:	Trawls through system to identify "the potential size of the problem".

Event card id: **Event 1.71**	
Actor:	Mr K
Action:	**Error explanation**
Event location:	Immunology department
Time/date event began:	24th May
Event duration:	
Data source:	10.34
Description:	Finds approx. 150 high-risk pregnancies which had been

	reported as low-risk.

Event card id: **Event 1.72**	
Actor:	Mr K
Action:	**Further investigation**
Event location:	Immunology department
Time/date event began:	11.30pm, 24th May
Event duration:	
Data source:	10.34
Description:	Emails findings [event 1.71] to Dr B.

Event card id: **Event 1.73**	
Actor:	Dr B
Action:	**Further investigation**
Event location:	Immunology department
Time/date event began:	11.30pm, 25th May
Event duration:	
Data source:	10.35
Description:	Sees email [event 1.72]

Event card id: **Event 1.74**	
Actor:	Mr J
Action:	**Further investigation**
Event location:	Immunology department
Time/date event began:	11.30pm, 24th May
Event duration:	
Data source:	10.34
Description:	Informs Chief Executive

A.2 Incident 2: Breast Cancer Screening errors, London

The data source for this incident is the formal Commission for Health Improvement report (CHI, 2002). The corresponding paragraph number will be entered in the source cell.

Event card id: **Event 2.1**	
Actor:	Patient
Action:	**No classification**
Event location:	
Time/date event began:	January 1999
Event duration:	
Data source:	1.1
Description:	Receives letter from WoLBSS indicating that her mammogram was normal.

Event card id: **Event 2.2**	
Actor:	Patient
Action:	**No classification**
Event location:	
Time/date event began:	2000 (before October)
Event duration:	
Data source:	1.1
Description:	Moves to area X

Event card id: **Event 2.3**	
Actor:	BSS X
Action:	**No classification**
Event location:	Breast Screening Service X
Time/date event began:	Mid-October 2000
Event duration:	
Data source:	1.1
Description:	Receives patient's file forwarded by WoLBSS

Event card id: **Event 2.4**	
Actor:	BSS X
Action:	**Error detection**
Event location:	BSS X
Time/date event began:	Mid-October
Event duration:	
Data source:	1.1
Description:	Realize patient has been screened incorrectly by WoLBSS by comparing their mammogram with the previous.

Event card id: **Event 2.5**	
Actor:	BSS X
Action:	**Error indication**
Event location:	BSS X
Time/date event began:	Mid-October
Event duration:	
Data source:	1.1
Description:	Contact senior manager at WoLBSS by phone regarding event 2.4.

Event card id: **Event 2.6**	
Actor:	BSS X
Action:	**Error indication**
Event location:	BSS X
Time/date event began:	31st October
Event duration:	
Data source:	1.2
Description:	Send written confirmation of the incident to WoLBSS.

Event card id: **Event 2.7**	
Actor:	Senior management
Action:	**Further investigation**
Event location:	WoLBSS
Time/date event began:	*1st November*
Event duration:	
Data source:	1.2
Description:	The matter of event 2.4 is discussed in management meeting.

Event card id: **Event 2.8**	
Actor:	Senior management
Action:	**Failure of Further investigation**
Event location:	WoLBSS
Time/date event began:	*1st November*
Event duration:	
Data source:	1.2
Description:	The matter of event 2.4 is not considered as important.

Event card id: **Event 2.9**	
Actor:	Quality assurance reference centre relating to BSS X
Action:	**Error indication**
Event location:	
Time/date event began:	1st or 2nd of November
Event duration:	
Data source:	1.3
Description:	Informs the London quality reference centre about event 2.4 ("verbally"—it is assumed this was done over the phone).

Event card id: **Event 2.10**	
Actor:	Quality assurance reference centre relating to BSS X
Action:	**Error indication**
Event location:	
Time/date event began:	Same day as event 2.9
Event duration:	
Data source:	1.3
Description:	Forward the letter that was sent to WoLBSS [event 2.4] to the London quality reference centre.

Event card id: **Event 2.11**	
Actor:	London quality reference centre
Action:	**No classification**
Event location:	London
Time/date event began:	10th November
Event duration:	
Data source:	1.3
Description:	Receive letter [event 2.10]

Event card id: **Event 2.12**	
Actor:	London quality reference centre
Action:	**Error indication**
Event location:	London
Time/date event began:	13th November
Event duration:	*5 days*
Data source:	1.4
Description:	Contacted senior management (WoLBSS) on "several occasions", insisting that the matter [event 2.4] be reported to the general manager of the imaging directorate of Hammersmith Hospitals NHS Trust)

Event card id: **Event 2.13**	
Actor:	WoLBSS
Action:	**Error indication**
Event location:	
Time/date event began:	17th November
Event duration:	
Data source:	1.5
Description:	Report the incident as critical incident to the general manager of the imaging directorate of Hammersmith Hospitals NHS Trust.

Event card id: **Event 2.14**	
Actor:	Trust chief operating officer/director of services
Action:	**Error indication**
Event location:	Hammersmith Hospitals Trust
Time/date event began:	17th November
Event duration:	
Data source:	1.6
Description:	Briefs the trust chief executive and medical director.

Event card id: **Event 2.15**	
Actor:	London quality reference centre
Action:	**Error indication**
Event location:	
Time/date event began:	22nd November
Event duration:	
Data source:	1.6
Description:	Informs the national coordinator of the NHS Breast Screening Programme and the chair of its administrative group.

Event card id: **Event 2.16**	
Actor:	London quality reference centre
Action:	**Error indication**
Event location:	
Time/date event began:	23rd November
Event duration:	
Data source:	1.6
Description:	Informs the officer with lead responsibility for cancer services at the NHS London regional office

Event card id: **Event 2.17**	
Actor:	London quality reference centre
Action:	**Error indication**
Event location:	
Time/date event began:	24th November
Event duration:	
Data source:	1.6
Description:	Confirms the incident in writing to the NHS London region director of public health.

Event card id: **Event 2.18**	
Actor:	London quality reference centre
Action:	**Further investigation**
Event location:	
Time/date event began:	4th December
Event duration:	
Data source:	1.7
Description:	Call a meeting to discuss the incident.

Event card id: **Event 2.19**	
Actor:	Meeting panel
Action:	**Further investigation**
Event location:	*Hammersmith Hospitals Trust*
Time/date event began:	4th December
Event duration:	
Data source:	1.7
Description:	An internal inquiry panel is established.

Event card id: **Event 2.20**	
Actor:	Inquiry panel
Action:	**Further investigation**
Event location:	*WoLBSS*
Time/date event began:	Early December
Event duration:	

Data source:	1.7
Description:	Review documents and conduct interviews.

Event card id: **Event 2.21**	
Actor:	Inquiry panel
Action:	**Error explanation**
Event location:	WoLBSS
Time/date event began:	Early December
Event duration:	
Data source:	1.7
Description:	Conclude that WoLBSS did not have a robust right results protocol.

Event card id: **Event 2.22**	
Actor:	Inquiry panel
Action:	**Further investigation**
Event location:	WoLBSS
Time/date event began:	Early December
Event duration:	
Data source:	1.7
Description:	Suggest that an external audit company reviews the mammogram files of all women who had attended for screening since 1993 (nearly 104,000 women and over 174,000 episodes)

Event card id: **Event 2.23**	
Actor:	Inquiry panel
Action:	**Further investigation**
Event location:	Hammersmith Hospitals Trust
Time/date event began:	21st December
Event duration:	
Data source:	1.8
Description:	Report their findings to a meeting.

Event card id: **Event 2.24**	
Actor:	Inquiry panel
Action:	**Further investigation**
Event location:	Hammersmith Hospitals Trust
Time/date event began:	22nd December
Event duration:	
Data source:	1.8
Description:	Report their findings to the trust chief executive.

Event card id: **Event 2.25**	
Actor:	PricewaterhouseCoopers
Action:	**Further investigation**
Event location:	WoLBSS
Time/date event began:	27th February, 2001
Event duration:	*Approx. 3 months*
Data source:	1.9
Description:	Go through the files of all women screened since 1993.

Event card id: **Event 2.26**	
Actor:	Inquiry panel
Action:	**Error Explanation**
Event location:	BSS X
Time/date event began:	July
Event duration:	
Data source:	1.10
Description:	Compile a report with their findings.

Event card id: **Event 2.27**	
Actor:	chief executive
Action:	**Error indication**
Event location:	Hammersmith Hospitals Trust's
Time/date event began:	9th March
Event duration:	

Data source:	1.11
Description:	Request CHI's assistance

Event card id: **Event 2.28**	
Actor:	CHI
Action:	**Further investigation**
Event location:	
Time/date event began:	10th April
Event duration:	
Data source:	1.11
Description:	Agree to conduct an investigation

Event card id: **Event 2.29**	
Actor:	WoLBSS
Action:	**Error control**
Event location:	
Time/date event began:	11th June
Event duration:	
Data source:	1.12
Description:	Suspend breast screening

Event card id: **Event 2.30**	
Actor:	CHI
Action:	**Further investigation**
Event location:	WoLBSS
Time/date event began:	11th June
Event duration:	*Approx. 3 months*
Data source:	1.11
Description:	Conduct investigation

Event card id: **Event 2.32**	
Actor:	WoLBSS
Action:	**Error correction**
Event location:	
Time/date event began:	10th December
Event duration:	
Data source:	1.12
Description:	Begin phased reintroduction of services.

A.3 Incident 3: Breast Cancer Screening errors, Manchester

Following this incident, two inquiries were conducted; the one by the Expert Advisory Panel [2006], and the other by Professor Mark Baker [2006]. The two reports have been used to produce the STEP cards for this incident. The EAP report has paragraph number, but not the Baker report. For the later, the page number will therefore be entered in the STEP card source cell.

Event card id: **Event 3.1**	
Actor:	Mammography radiographers
Action:	**Problem detection**
Event location:	Breast service, Trafford Hospitals Trust
Time/date event began:	April 2003
Event duration:	*8 months*
Data source:	Baker report, p6
Description:	Become concerned over Dr H's brusque style and high speed and brevity of reporting, the non-use of previous screening programme films in reporting mammograms and other practices, although no errors had been made.

Event card id: **Event 3.2**	
Actor:	Mammography radiographers
Action:	**Problem indication**
Event location:	Breast service, Trafford Hospitals Trust
Time/date event began:	November, 2003
Event duration:	
Data source:	Baker report, p6
Description:	Formally report their concerns [Event 3.1] to the *Trust management.*

Event card id: **Event 3.3**	
Actor:	THT management
Action:	**Further investigation**
Event location:	Breast service, Trafford Hospitals Trust
Time/date event began:	November, 2003
Event duration:	*<,unknown>, presumably less than a week*
Data source:	Baker report, p6
Description:	Investigate Dr H's practice.

Event card id: **Event 3.4**	
Actor:	THT management
Action:	**Failure of Further investigation**
Event location:	Breast service, Trafford Hospitals Trust
Time/date event began:	November, 2003
Event duration:	
Data source:	Baker report, p6
Description:	Do not find any errors or unacceptable behavior apart from the non-use of previous screening programme films.

Event card id: **Event 3.5**	
Actor:	Radiologists
Action:	**error detection**
Event location:	Trafford Hospital
Time/date event began:	November, 2003
Event duration:	
Data source:	Baker report, p6
Description:	Notice errors in Dr H's reports

Event card id: **Event 3.6**	
Actor:	Radiologists
Action:	**error indication**
Event location:	Trafford Hospital
Time/date event began:	November, 2003
Event duration:	
Data source:	Baker report, p6
Description:	Inform Trust management about Event 3.5

Event card id: **Event 3.7**	
Actor:	Trust management
Action:	**Failure of error recovery**
Event location:	Trafford Hospital
Time/date event began:	November, 2003
Event duration:	
Data source:	Baker report, p6

Description:	Consider errors as isolated events

Event card id: **Event 3.8**	
Actor:	Clinical staff
Action:	**problem detection**
Event location:	Trafford Hospital
Time/date event began:	November, 2003
Event duration:	
Data source:	Baker report, p6
Description:	Started becoming concerned that Dr H's work was not reliable.

Event card id: **Event 3.9**	
Actor:	Clinical staff
Action:	**Further investigation**
Event location:	Trafford Hospital
Time/date event began:	November, 2003
Event duration:	
Data source:	Baker report, p6
Description:	Would double-check Dr H's reports with other radiographers before communicating bad news to patients.

Event card id: **Event 3.10**	
Actor:	Mammography radiographers
Action:	**Error detection**
Event location:	THT
Time/date event began:	April, 2005
Event duration:	
Data source:	Baker report, p7
Description:	Find a high number of errors in a single MDT patient list.

Event card id: **Event 3.11**	
Actor:	Mammography radiographers
Action:	**Problem indication**
Event location:	THT
Time/date event began:	13th April, 2005
Event duration:	
Data source:	Baker report, p7, EAP report, 4.1
Description:	Report their concerns again to Trust senior management.

Event card id: **Event 3.12**	
Actor:	Mammography radiographers
Action:	**Error indication**
Event location:	THT
Time/date event began:	13th April, 2005
Event duration:	
Data source:	Baker report, p6, EAP report 4.1
Description:	Report event 3.7 to Trust senior management.

Event card id: **Event 3.13**	
Actor:	THT management
Action:	**Error control**
Event location:	Breast screening, THT
Time/date event began:	18th April, 2005
Event duration:	
Data source:	EAP report 4.1
Description:	Suspend Dr H

Event card id: **Event 3.14**	
Actor:	THT management
Action:	**Error indication**
Event location:	Breast screening, THT
Time/date event began:	19th April 2005
Event duration:	
Data source:	EAP report 4.1
Description:	report the issue as a Serious Adverse Event to the Greater Manchester Strategic Health Authority (GMSHA)

Event card id: **Event 3.15**	
Actor:	GMSHA
Action:	**Error indication**
Event location:	GMSHA
Time/date event began:	19th April 2005
Event duration:	
Data source:	EAP report 4.1
Description:	Inform the Department of Health.

Event card id: **Event 3.16**	
Actor:	Nightingale Centre
Action:	**Further investigation**
Event location:	Breast screening, THT
Time/date event began:	20 April 2005
Event duration:	
Data source:	EAP report 4.1
Description:	Are instructed to conduct an audit of Dr H's work

Event card id: **Event 3.17**	
Actor:	Nightingale Centre
Action:	**Further investigation**
Event location:	Breast screening, THT
Time/date event began:	26th April 2005
Event duration:	
Data source:	EAP report 4.1
Description:	Commencement of the review of the 457 Mammograms.

Event card id: **Event 3.18**	
Actor:	Expert Advisory Panel
Action:	**Further investigation**
Event location:	Breast screening, THT
Time/date event began:	26th April 2005
Event duration:	
Data source:	Baker report, p6
Description:	The panel is established

Event card id: **Event 3.19**	
Actor:	Nightingale Centre
Action:	**Error explanation**
Event location:	Breast screening, THT
Time/date event began:	6th May 2005
Event duration:	
Data source:	EAP report 4.1
Description:	Findings of initial mammography review (457 mammograms) highlighted a significant number of differing reports.

Event card id: **Event 3.20**	
Actor:	*THT management*
Action:	**Error indication**
Event location:	Breast screening, THT
Time/date event began:	27th May 2005
Event duration:	
Data source:	EAP report 4.1
Description:	Bury PCT became aware of the Serious Adverse Event due to a general alert regarding the excluded Consultant Radiologist being distributed to PCTs,

Event card id: **Event 3.21**	
Actor:	THT management
Action:	**Error indication**
Event location:	
Time/date event began:	17th June 2005
Event duration:	
Data source:	EAP report 4.1
Description:	NPSA notified of incident

Event card id: **Event 3.22**	
Actor:	Breast Screening Quality Assurance Team
Action:	**Further investigation**
Event location:	Breast screening, THT
Time/date event began:	1st August 2005
Event	*Approx. one month*

duration:	
Data source:	EAP report 4.1
Description:	External review of general radiology images/films

Event card id: **Event 3.23**	
Actor:	Breast Screening Quality Assurance Team
Action:	**Further investigation**
Event location:	Breast screening, THT
Time/date event began:	1st September 2005
Event duration:	*Apprx. 3 months*
Data source:	EAP report 4.1
Description:	External review of clinical notes, cytology reports and histopathology reports

Event card id: **Event 3.24**	
Actor:	Breast Screening Quality Assurance Team
Action:	**Further investigation**
Event location:	Breast screening, THT
Time/date event began:	1st September 2005
Event duration:	*Apprx. 4 months*
Data source:	EAP report 4.1
Description:	External review of ultrasound patients notes/images

Event card id: **Event 3.25**	
Actor:	Breast Screening Quality Assurance Team
Action:	**Error explanation**
Event location:	Breast screening, THT
Time/date event began:	21st November 2005
Event duration:	
Data source:	EAP report 4.1
Description:	Findings reported

Event card id: **Event 3.26**	
Actor:	Breast Screening Quality Assurance Team
Action:	**Further investigation**
Event location:	Breast screening, THT
Time/date event began:	1st December 2005
Event duration:	*One month*
Data source:	EAP report 4.1
Description:	External review of the 28 patients with delayed diagnosis completed

A.4 Incident 4: STI screening errors, Florida

As already mentioned, data for this analysis has been gathered with interviews. Therefore, the source cell will be left blank.

Event card id: **Event 4.1**	
Actor:	Biochemistry department
Action:	**No classification**
Event location:	Biochemistry department
Time/date event began:	January 2003
Event duration:	
Data source:	
Description:	New reporting system is introduced

Event card id: **Event 4.2**	
Actor:	Charge nurse
Action:	**No classification**
Event location:	Emergency Department
Time/date event began:	Early April
Event duration:	
Data source:	
Description:	Returns from maternity leave

Event card id: **Event 4.3**	
Actor:	Charge nurse
Action:	**Problem detection**
Event location:	Emergency Department
Time/date event began:	Early April
Event duration:	*2 weeks*
Data source:	
Description:	Notices there are no positives reported from Biochemistry department

Event card id: **Event 4.4**	
Actor:	Charge nurse
Action:	**Further investigation**
Event location:	Emergency Department
Time/date event began:	Mid April 2003
Event duration:	
Data source:	
Description:	Discusses matter (event 4.3) with a physician

Event card id: **Event 4.5**	
Actor:	Physician
Action:	**Further investigation**
Event location:	Emergency Department
Time/date event began:	Mid April
Event duration:	
Data source:	
Description:	Advises nurse to contact the lab

Event card id: **Event 4.6**	
Actor:	Charge nurse
Action:	**Problem indication**
Event location:	Emergency Department
Time/date event began:	Mid April 2003
Event duration:	
Data source:	
Description:	Informs lab that there are no positives being reported back to ED

Event card id: **Event 4.7**	
Actor:	Biochemistry department
Action:	**Further investigation**
Event location:	Biochemistry department
Time/date event began:	Mid April 2003
Event duration:	
Data source:	
Description:	Check tests carried out since change of equipment

Event card id: **Event 4.8**	
Actor:	Biochemistry department
Action:	**Error explanation**
Event location:	Biochemistry department
Time/date event began:	Mid April 2003
Event duration:	
Data source:	
Description:	Find positive tests in the system which have not been reported by reporting system

Event card id: **Event 4.9**	
Actor:	Biochemistry department
Action:	**Further investigation**
Event location:	Biochemistry department
Time/date event began:	Mid April 2003
Event duration:	
Data source:	
Description:	Check result reporting system

Event card id: **Event 4.10**	
Actor:	Biochemistry department
Action:	**Error explanation**
Event location:	Biochemistry department
Time/date event began:	Mid April 2003
Event duration:	
Data source:	
Description:	Identify/locate software bug

Event card id: **Event 4.11**	
Actor:	Biochemistry department
Action:	**Error correction**
Event location:	Biochemistry department
Time/date event began:	Mid April 2003
Event duration:	
Data source:	
Description:	Remove software bug

Event card id: **Event 4.12**	
Actor:	Biochemistry department
Action:	**Error correction**
Event location:	Biochemistry department
Time/date event began:	Mid April 2003
Event duration:	
Data source:	
Description:	Change printout settings so that all results are reported

Event card id: **Event 4.13**	
Actor:	Biochemistry department
Action:	**Further investigation**
Event location:	Biochemistry department
Time/date event began:	Mid April 2003

Event card id: **Event 4.16**	
Actor:	Charge nurse
Action:	**No classification**
Event location:	Emergency Department
Time/date event began:	Early week 3
Event	

Event duration:	
Data source:	
Description:	Search for all involved patients' details

Event card id: **Event 4.14**	
Actor:	Biochemistry department
Action:	**Error correction**
Event location:	Biochemistry department
Time/date event began:	End of April 2003
Event duration:	
Data source:	
Description:	Call patients back for screening

Event card id: **Event 4.15**	
Actor:	Charge nurse
Action:	**Problem detection**
Event location:	Emergency department
Time/date event began:	n/a, sometime in 2005
Event duration:	2 weeks
Data source:	
Description:	Notices there are no positives being reported back from the lab

duration:	
Data source:	
Description:	Recalls previous incident

Event card id: **Event 4.17**	
Actor:	Charge nurse
Action:	**Problem indication**

Event location:	Emergency Department
Time/date event began:	Early week 3
Event duration:	
Data source:	
Description:	Informs lab no positives are being reported back to ED (face-to-face)

Event card id: **Event 4.18**	
Actor:	Biochemistry department
Action:	**Further investigation**
Event location:	Biochemistry department
Time/date event began:	Early week 3
Event duration:	
Data source:	
Description:	Do not find any errors or unacceptable behavior apart from the non-use of previous screening programme films.

Event card id: **Event 4.19**	
Actor:	Biochemistry department
Action:	**Error explanation**
Event location:	Biochemistry department
Time/date event began:	Early week 3
Event duration:	
Data source:	
Description:	Realize a flat used to check whether a test result is positive has been mistakenly

	deactivated

Event card id: **Event 4.20**	
Actor:	Biochemistry department
Action:	**Error correction**
Event location:	Biochemistry department
Time/date event began:	Early week 3
Event duration:	
Data source:	
Description:	Reset flag

Event card id: **Event 4.21**	
Actor:	Biochemistry department
Action:	**Further investigation**
Event location:	Biochemistry department
Time/date event began:	Mid week 3, 2005
Event duration:	
Data source:	
Description:	Do not find any errors or unacceptable behavior apart from the non-use of previous screening programme films.

Event card id: **Event 4.22**	
Actor:	Biochemistry department
Action:	**Error correction**
Event location:	Biochemistry department
Time/date event began:	Mid week 3, 2005
Event	

duration:	
Data source:	
Description:	All involved patients are contacted for re-screening

Appendix B: ER-STEP exercise

This appendix will present the exercise that was used for the validation of the proposed technique, and the solutions provided by the two participants. The discussion regarding the results of validation activities can be found in Chapter 8:.

Focusing Accident and Incident Analysis on Error Handling: Error Recovery focused Sequentially Timed Events Plotting (ER-STEP).

Nick Chozos
PhD Student
Department of Computing Science
University of Glasgow
17, Lilybank Gardens
mailto:Nick@dcs.gla.ac.uk
07909793228

Accident and incident analysis techniques aim at identifying and understanding the factors that resulted to a disaster or a 'near- miss'. Such analyses are important, not only for the understanding of what caused an accident, but also for the understanding of how it could have been avoided, prevented, or better handled once the failure started to take place.

In the following pages, you will find a brief description of an incident analysis technique, and an exercise for its evaluation. The technique (ER- STEP) focuses on the analysis of error handling activities following the detection of a problem/ error/ technical fault. A brief incident scenario will be introduced, and participants will be asked to analyse the error handling efforts with the proposed analytical technique.

I would like to thank you for taking the time to conduct this exercise. Your findings will be compared to others' who have performed the same analysis, in order to evaluate how consistent and applicable the technique can be. If you have any queries please contact me at the email address above, or my PhD supervisor (Professor Chris Johnson, mailto:Johnson@dcs.gla.ac.uk).

ER- STEP: Focusing accident analysis on error handling activities

The technique presented here is an integration of an existing technique called Sequentially Timed Events Plotting (STEP), and the Error Recovery Framework, a theoretical framework that described the sequence of events that take place during error handling, from detection to recovery. STEP and the Error Recovery Framework will be introduced, before presenting Error Recovery focused- STEP, the integrated technique. An example will also be presented, in order to suggest how the technique should be applied during an incident analysis.

1. Sequentially Timed Events Plotting

STEP was developed by the USA Department of Energy. STEP is a reconstruction technique which presents the sequence of events as they evolved from left to right, denoting the actors' involvement. Figure 1 illustrates how STEP is applied in incident and accident analysis.

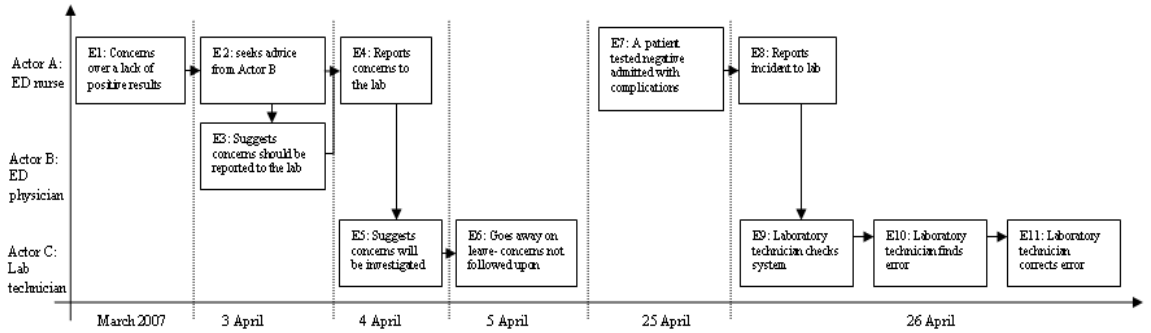


Figure 1: Example of incident analysis with STEP

In order to focus the analysis on error handling activities, STEP has been integrated with the Error Recovery Framework, which will be presented in the next section.

2. The Error Recovery Framework

The framework presented here is an adaptation of existing frameworks, while taking into account error detection theory from psychology and cognitive science. Figure 1 presents the sequence of the different stages that an error recovery process will go through. Each stage will be described in this section.

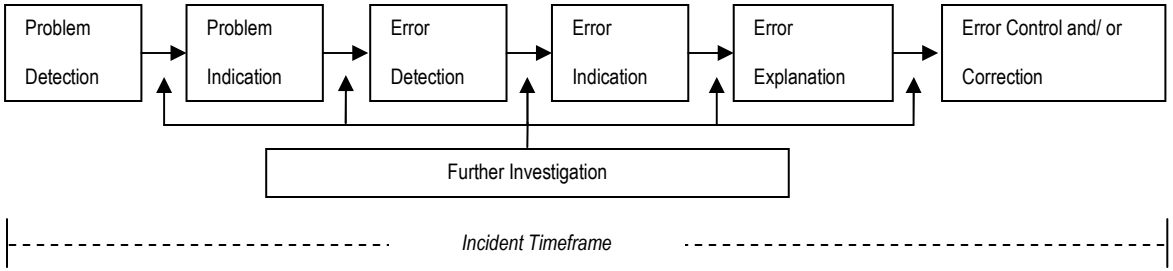


Figure 2: The Error Recovery Framework

Problem Detection: *Initial concerns that something may be wrong*. During *problem detection* the operator is not certain that there is an error or fault in the system. At this initial stage, the operator would have a 'feeling' based on experience and expertise, rather than any evidence suggesting concerns are valid. It is possible the operator might not act upon those concerns, but if they build up through time, the operator will either proceed with *problem indication*, or *further investigation*.

Error Detection: *A near miss, incident, or a system indication that something is wrong.* Unlike problem detection, error detection is a strong and clear indication of a system failure. However, error detection does not suggest the causes of the error have been identified. *Further investigation* might still be necessary in order to find the exact nature of the failure, and what caused it.

Further Investigation: *Actions to find out **if** there is an error, **what** the error is, what the **extent** of the failure is.* Activities range from seeking advice for further action to investigation into system elements that might be problematic, carrying out audits etc. Following *problem detection*, the operator might consult with a colleague, monitor system behaviour, or conduct other activities in order to determine if there is an error. *Further investigation* varies depending on the previous error recovery stage that took place. If it follows *error detection*, where more evidence is available, *further investigation* may include a range of different activities, all of which aim at finding out more about what caused the error.

Problem Indication: *Reporting of concerns (problem detection).* Reporting of concerns takes place when the operator that has detected a problem informs someone who can act upon these

concerns. It is important to distinguish *problem indication* from communication that takes place during *further investigation*. During *further investigation*, the person that detected a problem/ error, will communicate in order to find out more about the potentially problematic system element, while during *problem indication*, the operator will report his/ her concerns to someone who can do something about it.

Error Indication: *Reporting of error detection.* Error indication may take place in different forms: Incident report, communication over the phone, email etc.

Error Explanation: *The causes of the error are identified. Error explanation* usually follows *further investigation*. Once the causes of the failure have been identified and explained, *error correction* can be completed.

Error Correction: *Actions are taken to eliminate the error or fault that took place.* Modifications or interventions are introduced in order to assure the same error does not take place again.

The process of error recovery does not necessarily have to go through all stages. Also, some stages might be repeated, while the failure of a stage will either stop error recovery, or take it back to a previous stage.

3. Error Recovery- focused –STEP (ER STEP)

Figure 3 presents the application of ER STEP to the same sequence of events illustrated in figure 2. The classification of events has been done according to the different stages of the Error Recovery Framework.

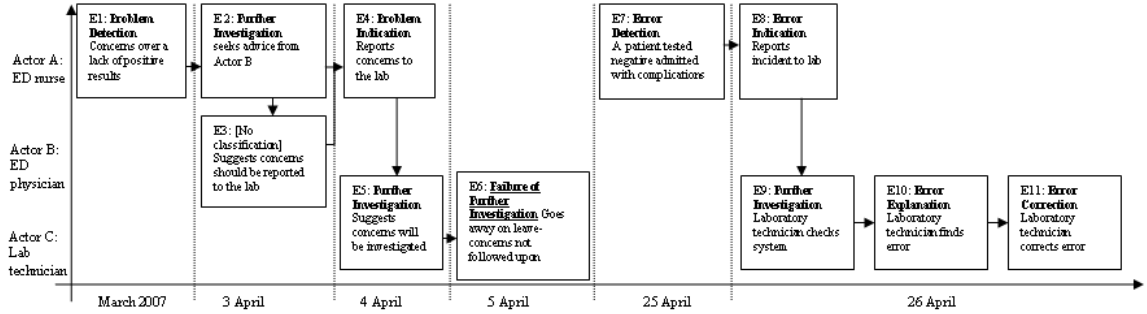


Figure 3: Analysis of error handling efforts with ER- STEP

Note event **E3** which is labelled as 'No Classification'. This is done when an event cannot be categorised under any of the stages of the error recovery process.

When the event following *further investigation* and/ or *problem/ error indication* results in slowing down or stopping the process, the event is labelled as 'Failure of- label of previous stage'. In the example, a Lab technician that was going to investigate the ED nurse's reports failed to do so. Event **E5** is then labelled as 'Failure of Further Investigation'.

The proposed technique aims at analysing error handling activities that took place during an incident. The labelling of events with the appropriate error recovery stage name, and the timely sequence can assist analysts to 'filter' activities, and to consider where error recovery was problematic.

However, in order for the technique to be applicable, it is important to have solid definitions of the error recovery stages, so that analyses are consistent and coherent. The purpose of this exercise is to evaluate this technique with its application to the same incident scenario by multiple participants. In the next section, the incident scenario will be presented.

4. Incident scenario

This scenario describes part of the events that took place during an incident in the Immunology Department of Sheffield Northern General Hospital. A software system used to calculate the likelihood of pregnant women giving birth to children with Downs Syndrome was not compliant with the millennium bug. As a result, calculations came back as negative, and pregnant women were screened as 'low- risk', even if they should be in the 'high- risk' area. The error resulted in the misdiagnosis of 235 women over a period of 5 months (January 1st to May 23rd 2000). From early on, several attempts were made by nurses that had concerns to contact the lab and resolve the matter. However, due to a number of reasons, error recovery efforts failed.

Following you will find 11 events that describe activities following January 1st 2000, during the response to the first indications of the software error.

4.1 Event Classification

Participants are asked to label the 11 events according to the stages of the Error Recovery Process, and then the sequence of events will be drawn in a STEP diagram as illustrated in figure 3. Events can be labeled as [*No Classification, Problem Detection, Error Detection, Further Investigation, Problem Indication, Error Indication, Error Explanation, Error Correction, Failure of Further Investigation, Failure of Problem Indication, Failure of Error Indication*]

Events following January 1st, 2000.

B.1 First participant's solution

E1: By the first two weeks of January, A Maternity and Gynaecology Liaison Sister at one of the hospitals (Hospital A) had become concerned as the number of screen positives she had received was lower than expected. _ *Problem Detection* _____

E2: Following her concerns, she made the Immunology Department aware of this at the end of January with a phone call, during which she spoke with Ms S, the acting MLSO. _____*Problem Indication* ____

E3: At the end of January, the Liaison Sister also queried a result for an older woman which she believed to be unrealistically low. She also spoke to Ms S about this matter.
_____*Problem Indication* _____

E4: Ms S had mentioned it to a colleague, Ms J in Hartlepool Hospital who had undertaken the Y2K test on January 4th. __ *Further Investigation* _____

E5: Ms J reassured her that everything was fine. __ *Failure of Further Investigation*_____

E6: Ms S did not log this incident because she did not regard it as very significant. _____ *Failure of Further Investigation* _

E7: The Liaison Sister decided to monitor her own screen positive results. _____ *Further Investigation* _____

E8: She phoned the Immunology Department again in April to express her concerns. She spoke to Mr M. _____*Problem Indication* _____

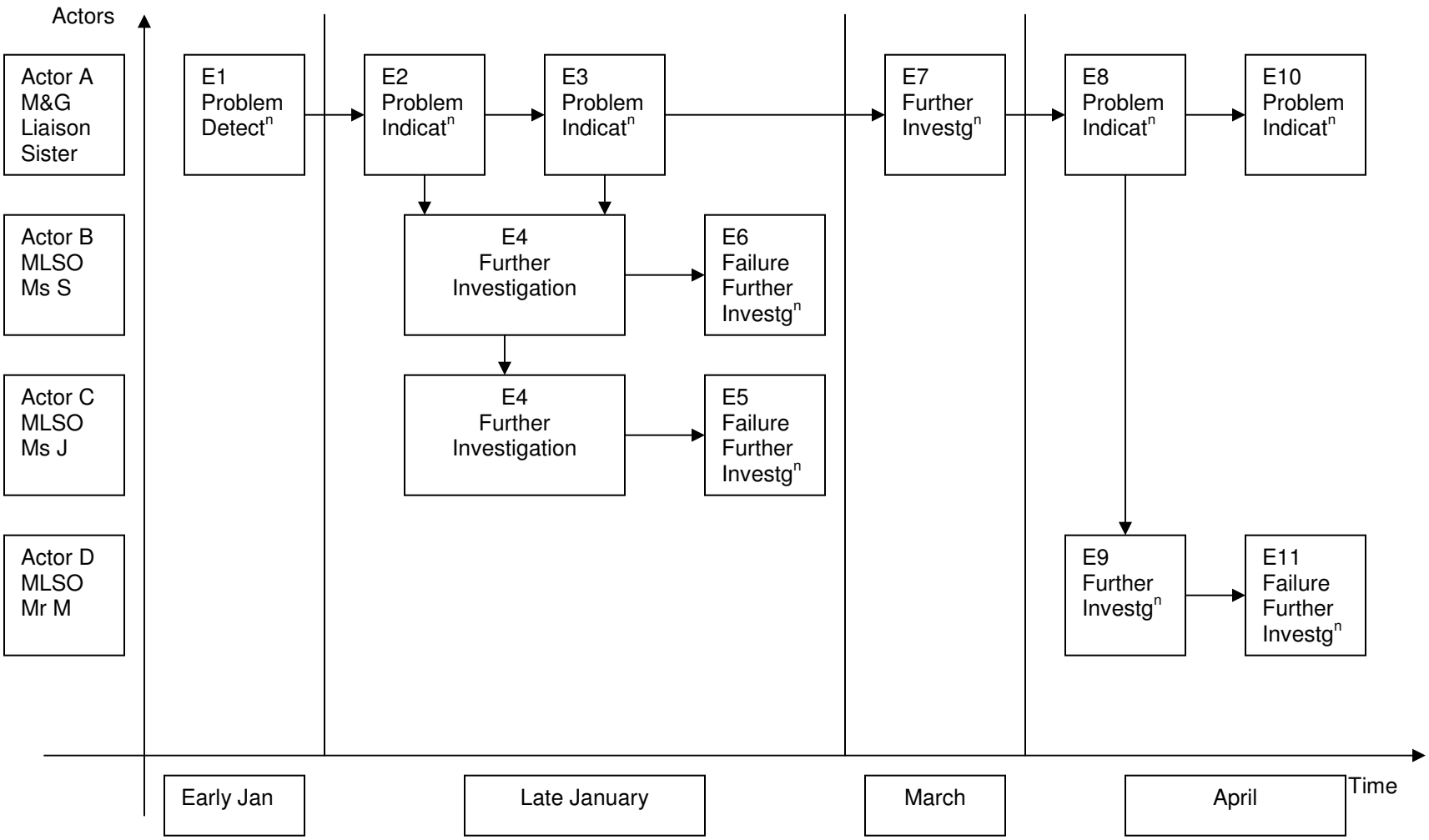
E9: Following her reports in April, she was told that it would be looked into by Mr M. _____*Further Investigation* _____

E10: During March and April, whenever results seemed unusual, the Liaison Sister would ring up and query the results for individual patients when they seemed unusual. _____ _____ *Problem Indication* _____

E11: On all occasions she received assurances that there was no problem. _____*Failure of Further Investigation* _____

4.2 ER STEP diagram

Now that events have been labeled, the ER STEP diagram can be drawn up. On the left of the vertical axis please place the title of each actor and the department/ organization they belong to. As events are drawn up from left to right, the corresponding date/ time of the event should be written below the horizontal axis, as in the example presented in Figure 3. Since events have been labeled in the previous part of the exercise, box diagrams do not need to contain the description of the event. It is advised that each box contains the event number and the error recovery stage name (e.g. E20: Error Detection)


Actors
Actor A
M&G
Liaison
Sister
Actor B
MLSO
Ms S
Actor C
MLSO
Ms J
Actor D
MLSO
Mr M
E1
Problem
Detect[n]
E2
Problem
Indicat[n]
E3
Problem
Indicat[n]
E4
Further
Investigation
E6
Failure
Further
Investg[n]
E4
Further
Investigation
E5
Failure
Further
Investg[n]
E7
Further
Investg[n]
E8
Problem
Indicat[n]
E10
Problem
Indicat[n]
E9
Further
Investg[n]
E11
Failure
Further
Investg[n]
Early Jan
Late January
March
April
Time

Please feel free to leave any comments in this page

Uncertain as to number of actual events – should they be broken down further?
Also uncertain as to time plotting - should timing of events be more detailed?
E.G. Problems are notified twice in separate events to Actor B Ms S, but event numbers do not detail this unless event label (E2) is repeated twice.
Same applies to E4..both Actor B Ms S and Actor C Ms J are engaged in 'further investigation'..separate events?
Events in March / April overlap…is loss of information on ER-STEP important?
Unclear whether or not Actor A's efforts end in failure.
Insufficient detail to report multiple 'problem indication' events from Actor A to unspecified persons, presumably in the lab who provided assurances.

B.2 Second participant's solution

E1: By the first two weeks of January, A Maternity and Gynaecology Liaison Sister at one of the hospitals (Hospital A) had become concerned as the number of screen positives she had received was lower than expected. __Problem Detection_____

E2: Following her concerns, she made the Immunology Department aware of this at the end of January with a phone call, during which she spoke with Ms S, the acting MLSO. _____Problem Indication_____

E3: At the end of January, the Liaison Sister also queried a result for an older woman which she believed to be unrealistically low. She also spoke to Ms S about this matter. _Problem Indication_____

E4: Ms S had mentioned it to a colleague, Ms J in Hartlepool Hospital who had undertaken the Y2K test on January 4th. __Further Investigation_____

E5: Ms J reassured her that everything was fine. ___Failure of Further Investigation ___

E6: Ms S did not log this incident because she did not regard it as very significant. _Failure of Problem Indication_____

E7: The Liaison Sister decided to monitor her own screen positive results. ____Further Investigation_____

E8: She phoned the Immunology Department again in April to express her concerns. She spoke to Mr M. ___Problem Indication____(? No definite proof of error)_____

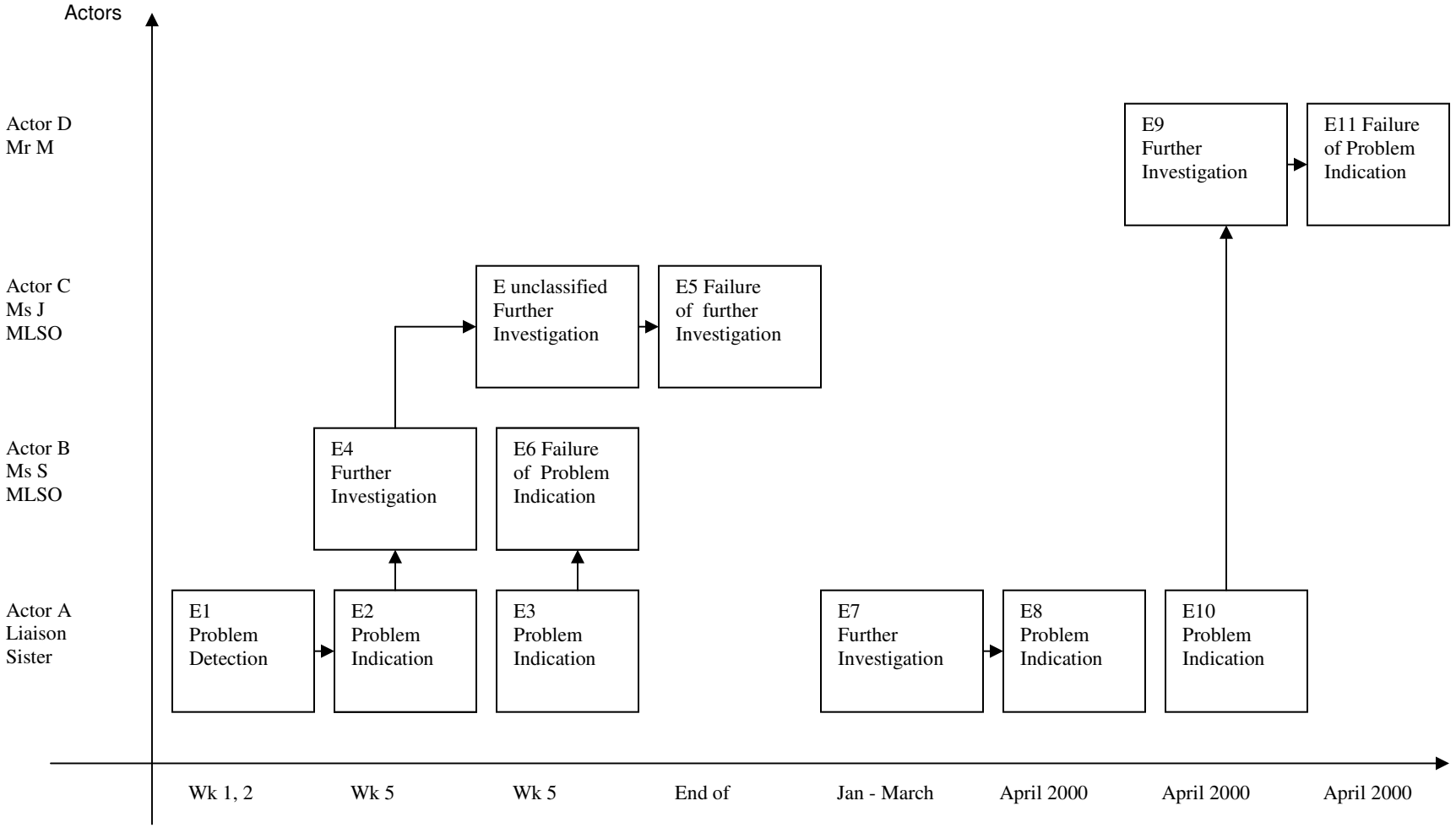
E9: Following her reports in April, she was told that it would be looked into by Mr M. ___Further Investigation_____

E10: During March and April, whenever results seemed unusual, the Liaison Sister would ring up and query the results for individual patients when they seemed unusual. ___Problem Indication_____

E11: On all occasions she received assurances that there was no problem. ___Failure of Problem Indication_____

4.3 ER STEP diagram

Now that events have been labeled, the ER STEP diagram can be drawn up. On the left of the vertical axis please place the title of each actor and the department/ organization they belong to. As events are drawn up from left to right, the corresponding date/ time of the event should be written below the horizontal axis, as in the example presented in Figure 3. Since events have been labeled in the previous part of the exercise, box diagrams do not need to contain the description of the event. It is advised that each box contains the event number and the error recovery stage name (e.g. E20: Error Detection)


Actors
Actor D
Mr M
Actor C
Ms J
MLSO
Actor B
Ms S
MLSO
Actor A
Liaison
Sister
E1
Problem
Detection
E2
Problem
Indication
E3
Problem
Indication
E4
Further
Investigation
E6 Failure
of Problem
Indication
E unclassified
Further
Investigation
E5 Failure
of further
Investigation
E7
Further
Investigation
E8
Problem
Indication
E10
Problem
Indication
E9
Further
Investigation
E11 Failure
of Problem
Indication
Wk 1, 2
Wk 5
Wk 5
End of
Jan - March
April 2000
April 2000
April 2000

Please feel free to leave any comments in this page

Found timing of events slightly difficult as not always discrete events, may cover a time period so become difficult to order and display

Could argue need to break down into more events but appreciate do not want to become swamped with detail. It may be better to have description of event in box.

Are repeated Problem Indications simply that or do they become a part of the Further Investigation process as they help to build a picture of the potential problem and start to confirm that it really exists.