

NICOLE / ISG
RISK ASSESSMENT COMPARISON STUDY

April 2004

Prepared for:

**Network for Industrially Contaminated Land in Europe (NICOLE)
Industrial Sub-Group (ISG)**

Prepared by:

**Geraghty & Miller International Inc.
2 Craven Court
Newmarket
CB8 7FA**

Tel: +44 (0) 1638 674767

Fax: +44 (0) 1638 668191

www.arcadisgmi.com

**RISK ASSESSMENT COMPARISON STUDY
NICOLE – INDUSTRY SUB-GROUP**

916830024_02 / April 2004

Client: NICOLE/ISG

Address: c/o Terry Walden, BP Global Environmental Management,
Sunbury-on-Thames, Middlesex, United Kingdom

Report Title: Risk Assessment Comparison Study

Report Number: 916830024

Report Date: April 2004

		Date	Signature
Report Authors	Elena Poletti Helen Hayward Jenny Gill Katy Baker		
Senior Review	Matt Gardner Cecilia MacLeod		
Authorisation	Lawrence Houlden		

If you have any questions regarding this project, please contact Matt Gardner.

TABLE OF CONTENTS

1.	INTRODUCTION	1
1.1.	Aims and Objectives.....	1
1.2.	Scope of Work	1
1.3.	Limitations.....	2
1.4.	Methodology.....	2
1.5.	Structure of the Report	3
2.	METHODOLOGY	4
2.1.	Phase I –Risk Assessment “System” Identification.....	4
2.2.	Phase II – System Screening.....	4
2.3.	Phase III – Generic Test Site (GTS).....	5
2.4.	Phase IV – Case Study Test Sites.....	6
3.	LITERATURE REVIEW	7
3.1.	Introduction	7
3.2.	Background.....	7
3.3.	Relevant Studies	7
3.3.1.	Ferguson and Kasamas (editors) (1999)	7
3.3.2.	Zaleski and Gephart (2000).....	8
3.3.3.	Whittaker <i>et al.</i> (2001).....	8
3.3.4.	Rikken <i>et al.</i> (2001).....	9
3.3.5.	Evans <i>et al.</i> (2002)	9
3.3.6.	Swartjes (2002)	10
3.4.	Conclusions	10
4.	REVIEW OF EUROPEAN RISK ASSESSMENT PRACTICE.....	11
4.1.	Introduction	11
4.2.	European Risk System Review.....	11
4.3.	Proprietary Systems	11
4.4.	Conclusions	12
5.	PHASE II – SYSTEM SCREENING.....	13
5.1.	Pro-forma Screening.....	13
5.2.	Matrix	13
5.3.	Criteria for System Screening.....	13
5.4.	Final System Selection	14
6.	PHASE III – GENERIC TEST SITE (GTS).....	16
6.1.	Introduction	16
6.2.	GTS Conceptual Site Model.....	17
6.2.1.	Sources	17
6.2.2.	Pathways.....	18
6.2.3.	Receptors.....	19
6.2.4.	Linkages in the GTS.....	19
6.3.	Results Comparison Methodology	19
6.3.1.	Receptor Point Concentration Comparisons	20
6.3.2.	Dose Concentration Comparisons.....	20
6.3.3.	Quantified Risk Level Comparisons	21
6.3.4.	Sensitivity Testing.....	22

**RISK ASSESSMENT COMPARISON STUDY
NICOLE – INDUSTRY SUB-GROUP**

916830024_02 / April 2004

6.4.	Input Parameter Selection.....	23
6.4.1.	Chemical and Toxicological Input Parameters	23
6.4.2.	Physical Parameter Selection	23
6.4.3.	Exposure Parameter Selection.....	24
7.	PHASE III – OUTPUT AND RESULTS.....	25
7.1.	Shallow Soil Exposure Pathways	25
7.1.1.	Soil Ingestion.....	25
7.1.2.	Dermal Contact	25
7.1.3.	Ingestion of Vegetables.....	26
7.1.4.	Inhalation of Dust Outdoors.....	27
7.2.	Deep Soil Environmental and Exposure Pathways	28
7.2.1.	Receptor Point Concentration in Indoor Air	29
7.2.2.	Dose Concentration from Inhalation of Indoor Air.....	30
7.2.3.	Groundwater Concentration at 50m from Deep Soil Source	31
7.2.4.	Surface Water Concentrations.....	32
7.2.5.	Surface Water Exposure - Dose Concentrations.....	33
7.3.	Groundwater Environmental and Exposure Pathways	34
7.3.1.	Receptor Point Concentration in Indoor Air	34
7.3.2.	Groundwater Concentrations at 50m Receptor Point.....	35
7.3.3.	Groundwater Ingestion Dose Concentrations	35
7.4.	Phase III Risk Calculations.....	36
7.4.1.	Results Presentation	36
7.4.2.	Toxicity	36
7.4.3.	Averaging Times Risk.....	37
7.4.4.	System Testing	38
8.	PHASE III SENSITIVITY TESTING.....	40
8.1.	Physical Input Parameters – Indoor Air.....	40
8.1.1.	Crack Fraction and Air Filled Porosity	40
8.1.2.	Pressure Difference	41
8.1.3.	Building Height	41
8.1.4.	Ventilation Rate	41
8.2.	Physical Input Parameters – Groundwater Migration	41
8.2.1.	Infiltration Rate	41
8.2.2.	Groundwater Flow Velocity.....	41
8.2.3.	Degradation Rate.....	42
8.3.	Chemical Input Parameters.....	42
8.3.1.	Dermal Contact with Soil.....	42
8.3.2.	Ingestion of Vegetables.....	43
8.3.3.	Indoor Air	43
8.3.4.	Groundwater (from Soil Leachate Source)	43
8.4.	Exposure Input Parameters	44
8.4.1.	Ingestion of Soil	44
8.4.2.	Inhalation of Indoor Air	44
8.4.3.	Ingestion of Vegetables.....	45
8.5.	Toxicological Input Parameters.....	45
9.	PHASE IV – CASE STUDY TEST SITES.....	47
9.1.	Introduction	47

**RISK ASSESSMENT COMPARISON STUDY
NICOLE – INDUSTRY SUB-GROUP**

916830024_02 / April 2004

10.	TEST SITE 1 - FORMER LUBE OIL PLANT	48
10.1.	Site Summary.....	48
10.2.	Background.....	48
10.3.	Conceptual Model.....	48
10.3.1.	Sources	48
10.3.2.	Pathways.....	49
10.3.3.	Receptors.....	50
10.3.4.	Input Data.....	50
10.3.5.	Output.....	50
10.3.6.	Scenarios	50
10.4.	Results and Discussion.....	51
10.4.1.	Receptor Point Concentrations.....	51
10.4.2.	Dose Prediction	52
11.	TEST SITE 2 - FORMER GASWORKS	53
11.1.	Summary.....	53
11.2.	Background.....	53
11.3.	Conceptual Model.....	53
11.3.1.	Sources	54
11.3.2.	Pathways.....	54
11.3.3.	Receptors.....	54
11.3.4.	Input Data.....	54
11.3.5.	Summary of Assumed Active Exposure Pathways	54
11.4.	Results and Discussion.....	54
12.	TEST SITE 3 - FLY ASH LANDFILL.....	58
12.1.	Summary.....	58
12.2.	Conceptual Model.....	59
12.2.1.	Sources	59
12.2.2.	Pathways.....	59
12.2.3.	Receptors.....	59
12.2.4.	Input Data.....	60
12.2.5.	Summary of Assumed Active Exposure Pathways	60
12.3.	Results and Discussion.....	60
12.3.1.	Ingestion of soil.....	60
12.3.2.	Dermal Contact with Soil.....	61
12.3.3.	Ingestion of vegetables.....	61
12.3.4.	Inhalation of dust outdoor	62
13.	TEST SITE 4 - ACTIVE CHEMICAL MANUFACTURING SITE	63
13.1.	Introduction	63
13.2.	Conceptual Model.....	63
13.2.1.	Sources	63
13.2.2.	Pathways.....	64
13.2.3.	Receptors.....	64
13.2.4.	Input Data.....	64
13.2.5.	Summary of Assumed Active Exposure Pathways	64
13.3.	Results and Discussion.....	64
13.3.1.	Receptor Point Concentrations.....	64
13.3.2.	Doses	65

**RISK ASSESSMENT COMPARISON STUDY
NICOLE – INDUSTRY SUB-GROUP**

916830024_02 / April 2004

14.	TEST SITE 5 - FORMER RETAIL PETROL FILLING STATION	67
14.1.	Introduction	67
14.2.	Conceptual Model.....	67
14.2.1.	Sources	67
14.2.2.	Pathways.....	67
14.2.3.	Receptors.....	67
14.2.4.	Input Data	68
14.2.5.	Summary.....	68
14.3.	Results and Discussion	68
14.3.1.	Receptor Point Concentrations.....	68
14.3.2.	Doses	69
15.	CONCLUSIONS AND RECOMMENDATIONS	70
15.1.	Conclusions	70
15.2.	Recommendations.....	73
16.	REFERENCES	74

Tables

1.	Summary of Risk System Capabilities
2.	Risk System versus Screening Criteria Matrix
3.	Chemical and Toxicological Properties
4.	Risk System versus Source-Pathway-Receptor Links Matrix
5.	Key Physical Input Parameters for the Generic Test Site
6a.	Key Exposure Parameters for the Generic Test Site - Software Symbols
6b.	Key Exposure Parameters for the Generic Test Site - Defaults for Adult Receptor
6c.	Key Exposure Parameters for the Generic Test Site - Defaults for Child Receptor
7a.	Risk Levels for TCE for Child Receptor
7b.	Risk Levels for TCE for Adult Receptor
8.	Risk Levels for Benzene for Child and Adult Receptors
9.	Sensitivity Analysis: Physical Input Parameters
10.	Sensitivity Analysis: Chemical Property Values
11.	Sensitivity Analysis: Exposure Input Parameters
12.	Sensitivity Analysis: Toxicological Input Parameters
13a.	Risk Levels for TCE: Child Receptor Based on GTS
13b.	Risk Levels for TCE: Child Receptor Based on Default Toxicity Data
13c.	Risk Levels for TCE: Adult Receptor Based on GTS
13d.	Risk Levels for TCE: Adult Receptor Based on Default Toxicity Data
14a.	Risk Levels for Benzene: Child & Adult Receptor Based on GTS
14b.	Risk Levels for Benzene: Child & Adult Receptor Based on Default Toxicity Data
15.	Summary of Pathways Considered for Test Sites
16.	Risk System vs. Test Site Exposure Pathways Matrix
17.	Key Physical Input Parameters – Test Site One
18.	Results – Test Site One
19.	Key Physical Input Parameters – Test Site Two
20.	Results – Test Site Two
21.	Key Physical Input Parameters – Test Site Three
22.	Results – Test Site Three
23.	Key Physical Input Parameters – Test Site Four
24.	Results – Test Site Four
25.	Key Physical Input Parameters – Test Site Five
26.	Results – Test Site Five

Figures

1. Methodology Flow Diagram
2. Source-Pathway-Receptor Links
3. GTS – Schematic Conceptualisation in Cross Section
4. Ingestion of Soil Dermal Contact Pathways – Adult Doses
5. Vegetable Ingestion from Soil – Adult Doses
6. Inhalation of Dust Outdoors – Adult Doses
7. Indoor Air Concentrations from Soil Source
8. Flow Diagrams for Soil to Indoor Air Models
9. Inhalation of Indoor Air from a Soil Source – Adult Doses
10. Groundwater Concentrations from Soil Source
11. Flow Diagrams for Soil to Groundwater Migration Models
12. Surface Water Concentrations from Soil Source
13. Flow Diagrams for Soil to Surface Water Models
14. Surface Water Exposures – Adult Doses
15. Indoor Air Concentrations from Groundwater Source
16. Flow Diagrams for Groundwater to Indoor Air Models
17. Groundwater Concentration from Groundwater Source
18. Flow Diagrams for Groundwater Fate and Transport Models
19. Ingestion of Groundwater – Adult Doses
- 20.-23. Sensitivity Testing of Physical Input Parameters
- 24.-25. Sensitivity Testing of Chemical Input Parameters
26. Sensitivity Testing of Exposure Input Parameters
27. Site Conceptual Model – Test Site One
28. Results – Test Site One
29. Site Conceptual Model – Test Site Two
- 30.A-D Results – Test Site Two
31. Site Conceptual Model – Test Site Three
- 32.A-D Results – Test Site Three
33. Site Conceptual Model – Test Site Four
- 34.A-D Results – Test Site Four
35. Site Conceptual Model – Test Site Five
- 36.A-B Results – Test Site Two

Appendices

- A. Abbreviations and Glossary
- B. European Contaminated Land Risk Assessment Review
- C. Blank Pro-forma Questionnaire
- D. Completed Risk System Questionnaires
- E. Phase III Risk Systems Review
- F. Specific Model Parameter Inputs

1. INTRODUCTION

In January 2001, Geraghty and Miller International Inc (“Arcadis GMI”) was approached by Terry Walden of BP Global Environmental Management to submit a proposal to the Industrial Sub-Group (ISG) of the Network for Industrially Contaminated Land in Europe (NICOLE) to undertake a comparative study of contaminated land risk assessment tools. A proposal was submitted and interest was sufficient from the NICOLE-ISG that the project was initiated.

The Client for this project is therefore described as NICOLE/ISG. Eleven members of the ISG have provided the funding for this project and NICOLE has provided additional funding. A list of participating companies and organisation is presented below:

Akzo Nobel	NICOLE
BNFL	PowerGen
BP	SecondSite Property
Fortum	Shell
ICI	Solvay
JM Bostad	TotalFinaElf

1.1. Aims and Objectives

“The aim of this study is to critically appraise the human health risk assessment models/systems commonly used in the different countries of Europe.”

The objectives behind this aim are as follows:

- Increase awareness and understanding of model variability for Risk Assessors applying the models.
- Provide confidence to Risk Managers relying on model output to facilitate environmental decision-making.

1.2. Scope of Work

The focus of this study is on site-specific human health risk assessment and therefore it should be noted that other forms of environmental risk assessment, such as ecological risk assessment or building risk assessment are not evaluated.

The comparison of risk assessment systems will include the following:

Phase I and Phase II

The identified contaminated land risk assessment systems are reviewed and compared with respect to capability to assess pathways and with respect to outputs.

Phase III and Phase IV

Each risk assessment system selected for the generic test site and the case study test sites is used to assess the defined source-pathway-receptor linkages. Results for Receptor Point concentrations, Dose Concentrations and Quantifiable Risk Levels are compared and conclusions are drawn as to where similarities and differences exist. It must be stressed that risk assessment systems are not ranked in relative superiority, and it is not an aim of the study to distinguish some systems as being better or worse than others.

1.3. Limitations

Each system has been discussed within the remit of the project resulting in its development, taking into account existing or pending legislative requirements together with consideration of the socio-economic status of the origin state. For example, a project to develop a risk assessment package in the UK may not have had the same aims or objectives as a project to develop a risk assessment package in Italy or the Netherlands. Therefore a direct comparison between the systems as a whole would be unjustifiable. This project aims to make meaningful comparisons between systems that purport to provide a similar capability. To this end not all of the systems selected can be compared for every factor considered.

In undertaking this comparison study, Arcadis GMI has only used the version of software packages that were available at the time. We have not obtained access to any of the packages at the software coding level and are using the systems as they exist on the market.

Model validation and verification are critical processes in software development, which should be undertaken prior to the release of modelling tools such as those used within risk assessment systems. Therefore, it is not within the scope of this study to carry out such validation work. It has been assumed that the models used within this comparison study have been independently validated prior to this work.

Some of the risk assessment systems incorporate a stochastic or probabilistic modelling capability. These features of those systems are not tested, except where a deterministic capability does not exist, however attention is drawn to the capability existing.

1.4. Methodology

To complete the project, a phased methodology has been devised, comprising:

- Phase I Identification of Risk Assessment “Systems” in European countries;
- Phase II Pro-forma screening of Systems to evaluate capabilities;
- Phase III Use of a hypothetical Generic Test Site data set to produce comparable results;
- Phase IV Use of 5 Case Study Test Sites to evaluate scenarios in non-idealised situations.

The methodology is described in greater detail in Section 2.

1.5. Structure of the Report

The following list outlines the structure of this report. Each point being expanded further in the report:

1. Introduction
2. Methodology
3. Phase I - Literature Review
4. Phase I - European Contaminated Land Risk Assessment Review
5. Phase II - Screening of Risk Systems
6. Phase III - Generic Test Site Definition and Simulations
7. Phase III – Output and Results
8. Phase III - Sensitivity Testing
- 9.-14. Phase IV - Case Study Test Sites Definition and Simulation
15. Conclusions and Recommendations
16. References

Contaminated land risk assessment is a highly specialised and technical subject and as such it has a “technical language”. Add to this the language differences between the nations that have developed risk assessment systems and it can be seen that an existing uniform description is unlikely. Therefore a glossary of terminology is included as Appendix A.

2. METHODOLOGY

A phased approach has been developed that commences with an initial examination of risk assessment “practice” across Europe including comment on legislative tools, generic guidelines or the existence of site specific risk assessment guidelines. This is followed by a series of systematic reviews and screens to focus the level of comparison from generic elements that are more widespread down to site-specific elements which are fewer.

The structure of the project allows the examination and comparison of the maximum amount of information whilst identifying a limited number of risk assessment systems for more intensive examination at subsequent Phases.

Figure 1 illustrates the process adopted and a description of the process is provided below.

2.1. Phase I – Risk Assessment “System” Identification

This phase comprises two elements.

- **Literature Review:** The first element is a review of existing literature that overviews contaminated land risk assessment in Europe. Several previous projects have been undertaken with similar aims and objectives to this project, a review of these studies led to the development of the methodology. Furthermore, this review has directed the project away from repeating previous studies while allowing us to target questions raised in these reports.
- **European Contaminated Land Risk Assessment Review:** The second element of the review comprised an examination of the legislative management of contaminated land in European countries with the aim of identifying the appropriate means of assessing potential risks in each country. The conclusion of this element is the determination of a list of risk assessment “systems” for consideration at Phase II.

2.2. Phase II – System Screening

The list of risk assessment systems generated from Phase I include methodologies, guidance documents, screening tools and software tools all of which assist in decision making. In Phase II, each system is evaluated against a pro-forma questionnaire to evaluate a system’s capabilities. The pro-forma questions the functionality of each system in terms of capability for example what types of exposure pathways it considers, whether it incorporates fate and transport models and what outputs are produced. The conclusion of Phase II is a list of systems suitable for further quantitative testing at Phase III based on the following criteria:

1. Can the system be used for site specific risk assessment?
2. Can the system quantify human health risk?
3. Is this the sole system developed in the country of origin?

2.3. Phase III – Generic Test Site (GTS)

This Phase of the project tests the capabilities of each system against a generic data-set with the aim of identifying similarities and differences between:

- Data entry requirements
- Environmental fate and transport algorithms
- Human health exposure algorithms
- Risk calculation equations

Phase III comprises the development of a Conceptual Site Model (CSM), the development of a consistent data set between the relevant systems, the assessment of the source pathway receptor linkages in the CSM using the applicable risk assessment systems and the testing of sensitivity in outputs to changes in parameter inputs.

The GTS CSM is a set of source-pathway-receptor scenarios defined to represent typical human health exposure pathways at any impacted site. The GTS model then requires defining in terms of input parameters. For several of the exposure scenarios, the data entry requirements are different and checks must be made to ensure that the input parameters are consistent between systems and are in agreement with the hypothetical exposure scenario.

Each scenario is assessed with the risk assessment systems that can complete an assessment for that pathway. The output from the model runs are collated and reviewed to allow comparison between results. Differences in output can be attributed to differences between how each system conceptualises the exposure scenario, what algorithm or set of algorithms the system uses to simulate the pathway or where an assumption is built into a system (hard-wired parameters).

To facilitate a greater understanding of the similarities and variations between different systems, the sensitivity of major parameters is tested. It is not within the scope of this study to test the sensitivity of every parameter; therefore some key parameters are tested including groups of parameters where it is the sensitivity of the group that is important rather than the individual parameters. The methodology of the analysis is detailed below:

- Test sensitivity to environmental parameter inputs, the following key parameter sets are selected for testing:
 - Vapour intrusion rates
 - Groundwater velocity
- Test sensitivity to chemical property inputs, all risk assessment system chemical specific defaults are re-applied and the output is compared with the results of the GTS assessments.
- Test sensitivity to exposure property inputs, all risk assessment system exposure specific defaults are re-applied and the output is compared with the results of the GTS assessments.
- Test sensitivity to toxicological property inputs, all risk assessment system toxicological specific defaults are re-applied and the output is compared with the results of the GTS assessments.

A range of parameter values is defined for each selected parameter using the generic dataset as the ‘start value’ for each system. Each parameter or group of parameters are modified, one at a time, whilst maintaining the remaining parameters at the start value. The difference in output is compared to the difference in input and conclusions drawn.

2.4. Phase IV – Case Study Test Sites

The aim of Phase IV is to draw out and expand upon the conclusions made at Phase III using real site data sets. Many of the conclusions drawn from Phase III can be attributed to the simplifications made in developing the GTS data set.

Possibly the greatest level of uncertainty in risk assessment is created when conceptualising the environment. The assignment of physical parameter values based on what is often limited site data results in high levels of subjectivity and reliance on conservatism/default parameters. In conceptualising the test sites, Arcadis has produced what we believe to be reasonable conceptual models based on experience, however a different conceptual model could be determined with additional data or greater levels of local awareness and regulatory control. Where site-specific data does not exist, Arcadis has deferred to the default input parameters used in each of the systems. If a model does not have a default value, Arcadis has adopted values based on literature, or drawn from default values used in other model/systems.

Chemical Parameters

Chemical specific values have been taken as the defaults in each system and have not been standardised across the systems. Where a chemical was not included in the database, parameters have been adopted based on defaults used in other systems or literature values.

Toxicological Parameters

Toxicological values have been taken as the defaults in each system and have not been standardised across the models.

Exposure Parameter

Where possible, if there is no site specific data, exposure data is taken to be the defaults in the system databases. If more than one default dataset is available, one has been selected and the effects of selecting one over the other is discussed.

Five Case Study Test Sites have been defined for assessment:

1. Former Lube Oil Plant
2. Former Gasworks Site
3. Fly Ash Landfill
4. Active Chemicals Manufacturing Site
5. Former Retail Petrol Filling Station

Each of the sites is assessed for a range of applicable source-pathway-receptor linkages. Not all exposure pathways are active at any one site, therefore a range of different types of site and different types of contaminants of concern have been selected in order to explore as many of the previously investigated pathways as possible. The sites selected represent “typical” sites, but they have also been selected on the grounds that they comply with a number of the assumptions adopted for the environmental fate and transport models. The sites are relatively easily conceptualised, but include natural variations that may potentially result in larger variations in results. Furthermore the sites have been selected on the grounds that they have good data coverage, which may often be missing at similar sites.

3. LITERATURE REVIEW

3.1. Introduction

Risk assessment is widely recognised as a key concept in the management of contaminated land, however the uptake of use by contaminated land investigators and the tools available to assist in the assessment are at differing levels of development in the countries of Europe. Furthermore, where risk assessment systems exist, there is often one or more ways of completing the necessary tasks.

The aim of this section is to review existing literature relevant to this project in terms of aims, objectives, methodology of study and conclusions drawn. The review is not exhaustive, but has singled out several key reports from which valuable insight has been obtained.

3.2. Background

The adoption of “risk assessment” as a means for assessing and managing the potential impacts associated with contaminated land began in the early 1970’s. It is generally accepted that there are two main centres where this pragmatic management tool is rooted and now well developed; the Netherlands and the United States of America.

A systematic and defensible framework for the assessment of contaminated land was developed in the 1970’s for the US Superfund, whereas the Dutch Intervention Values are widely used as preliminary screening values for soil and groundwater throughout Europe.

The underlying concepts, at least, have been adopted in the remaining countries of Europe, however they have been developed to differing levels. Different national frameworks have produced environmental risk assessment systems that reflect the varying legislative requirements and also the socio-economic and environmental diversity of each country.

3.3. Relevant Studies

3.3.1. Ferguson and Kasamas (editors) (1999)

The Concerted Action on Risk Assessment for Contaminated Sites in Europe (CARACAS) commissioned a two-volume desk reference book addressing the aims of risk assessment and the variety between risk assessment practises in the 16 CARACAS countries.

Volume 1 commences with a review of the fundamental concepts of risk assessment. Subsequent chapters cover the various technical aspects of risk assessment in greater detail. Volume 1 also introduces the source-pathway-receptor conceptual model, the differences between generic and site-specific risk assessment, the differences between human health and environmental risk assessment and attenuation mechanisms.

Volume 2 presents a commentary on the status and evolution of risk assessment practice in each of the 16 countries affiliated with CARACAS. Each nation’s policy and practice is described from legislation to the development of tools to assist in the assessment and management of contaminated land. The appendices to Volume 2 present some useful comparative matrices of risk assessment use in separate CARACAS countries.

The final section of Volume 2 of the Caracas books introduces knowledge sharing organisations including CARACAS itself, but also the Network for Industrially Contaminated Land in Europe (NICOLE) and the Contaminated Land Rehabilitation Network for Environmental Technologies (CLARINET).

3.3.2. Zaleski and Gephart (2000)

In recognition of the need to use site specific data in exposure assessment modelling, NICOLE commissioned the development of the Exposure Factors Sourcebook. Human “lifestyle” factors or exposure factors are needed to calculate predicted maximum doses for risk assessment, and in defining acceptable doses in remedial target derivation. Therefore the factors are fundamental to human health risk assessment.

Whilst this publication is not a study between risk assessment software tools it contains valuable information on the variability between the populations of different countries and highlights the necessity to recognise that site specific risk assessment incorporates nation/region specific data. The book compiles and presents all the relevant human health exposure factors data available from national and international studies.

3.3.3. Whittaker *et al.* (2001)

A research project was commissioned by the Environment Agency of England and Wales (EA), with the aim of producing a methodology for benchmarking models for use in conjunction with the EA’s *Methodology for the Derivation of Remedial Targets for Soil and Groundwater to Protect Water Resources* (R&D Publication 20 – Marsland and Carey, 1999). Two software tools (a worksheet accompanying the R&D P20 report and *ConSim*) were developed for the EA to implement the P20 methodology. However it was recognised that alternative tools were also in use across the UK and therefore a selection of four software tools were included: Remedial Targets Worksheet v1.1 (RTW), ConSim v1.05, RBCA Toolkit for Chemical Release v1.2 and RISC v.3.09. To assist in the benchmarking process, four case studies were used as test sites to assess performance and allow comparison of results.

The methodology followed in the study is presented below:

- A pro-forma system was developed, whereby the software tool was described in terms of functionality, pathways assessed, sources and receptors available and inputs and outputs.
- Case studies were developed to test the software tools within the context of groundwater risk assessment following the P20 Methodology. Each scenario was described, then the case study was worked through in stages, which corresponded to tiers in some of the software packages, or sub-divided tiers in others.
- Remedial target concentrations were reported in some cases, and receptor point concentrations for the remaining case studies. Results were recorded to three significant figures, which was deemed sufficient to highlight significant differences in the modelling approach.
- Where differences were highlighted, an effort was made to explain the origin of the difference, to suggest methods of ‘correcting’ the parameters or adapting model input and to applying such new parameters to reproduce the results of the other systems.

The report concluded that the four software tools have a similar approach to modelling groundwater flow and contaminant transport, while a number of significant differences were also highlighted, including:

- 1) Three of the four systems applied first-order decay to both the dissolved and sorbed phases, leading to a higher degradation rate than for dissolved phase degradation alone; and
- 2) Where the leachate flux is not insignificant with respect to the receiving groundwater flux, errors in contaminant travel times may arise for RBCA Toolkit and the RTW.

The results obtained from the four software tools were within an order of magnitude or so, and none was deemed consistently more or less conservative.

3.3.4. Rikken *et al.* (2001)

The aim of this study was to evaluate the models used in the CSOIL human exposure model by comparing with a group of alternative human exposure models. The CSOIL model had been used in the Netherlands to develop the Dutch Intervention Values (DIVs) since 1991.

The study begins with a comparison of CSOIL with the European Union System for the Evaluation of Substances (EUSES), this is followed by comparison with the Contaminated Land Exposure Assessment (CLEA) system from the UK, the Umwelt Menschen Schadstoffe (UMS) system from Germany and the US-EPA CalTox system. Comparison is limited to assessing risks from soil impacts and because of the aim to compare with CSOIL, the pathways are limited to those included in CSOIL however, specific focus is given to the modelling of volatiles in indoor air and metal uptake in plants.

3.3.5. Evans *et al.* (2002)

A research project was commissioned by the Environment Agency of England and Wales (EA), and contracted to Golder Associates (UK) Ltd, with the aim of evaluating soil vapour intrusion models for use in the Contaminated Land Exposure Assessment (CLEA) system. The systems selected for review exist as either separate stand alone models or are incorporated into existing risk assessment software programs.

One of the key objectives of the study was to undertake a review of the characteristics of the identified soil vapour models. The methodology followed is presented below:

- Existing soil vapour intrusion systems were critically reviewed, in terms of processes simulated, the rigour of the fate and transport algorithms, the functionality of the software system during the inputting of data and the commercial availability of the system.
- The most suitable models were selected for a sensitivity analysis, based on apparent benefits and limitations of the models.
- Published and unpublished soil vapour intrusion data sources (case study test sites) were reviewed and the most suitable data-sets chosen for use in the calibration process of the models.
- A sensitivity analysis was undertaken, using a generic data-set and the selected models.
- The selected models were calibrated against the case study test data.

Ten systems were included in the initial screening with a shortlist of five chosen for further assessment against the case study site data. Of the five systems compared, the system deemed most suitable for inclusion in the CLEA system was the Johnson and Ettinger sub-model (as used in the RISC v3.09 package).

3.3.6. Swartjes (2002)

The study aimed to compare the dose calculations of a variety of contaminants for generic scenarios, using risk assessment systems from seven European countries (Denmark, UK, The Netherlands, France, Sweden, Italy, Belgium/Flanders).

Three aims of the comparative study were identified:

- 1) Investigate variations in calculated human exposure level for different systems;
- 2) Investigate variation in default input parameters for different systems; and
- 3) Evaluating differences in calculated exposure via different major exposure routes on the basis of differences between system concepts and input parameters.

The methodology used commenced with a pro-forma review of system capabilities, which was sent to and filled in by the “owners” of the risk assessment system, i.e. the regulatory authority for that country. Following this, system users in each of the seven countries produced calculated exposure levels for the different exposure pathways via the appropriate risk assessment system. Both a generic data set of input parameters and a system-specific (default) data set of input parameters were tested. To demonstrate variability, results were plotted as variations from the common median for each output of concern. An overview of the system-specific default values was completed for the main input parameters and a qualitative evaluation of differences between the system outputs was completed.

A wide range of system calculation results was reported. The greatest degree of variability was reported where fate and transport models were incorporated into the assessment, particularly in the case of vapour migration. It was outside the scope of the project to fully explain all the variations observed in terms of differences in total exposure or specific exposure pathways with respect to system concepts.

3.4. Conclusions

The comparison study presented within this report aims to build upon the results of the research discussed above, however it is not plausible to cover all the issues presented. The focus of this study remains the comparison of human health risk assessment systems, however an explanation of variability between systems is also desired and therefore examination of each system at the algorithm level has been undertaken together with the examination of the use of selected default parameters.

The methodology used in this study is in general agreement with that of Whitaker *et al.* (2001) in model selection and review. The methodology also incorporates that of Swartjes (2002), including the evaluation and comparison of model functionality for a generic data set. However, a difference exists in that each system is reviewed by an independent third party and not the “system-owners”.

In addition to the generic test site data, this study uses ground data from five case study sites. This approach was used by Whittaker *et al.* (2001) and Evans *et al.* (2002) and serves the purpose of testing systems against non-idealised data.

4. REVIEW OF EUROPEAN RISK ASSESSMENT PRACTICE

4.1. Introduction

This section compliments Section 3.0 in completing Phase I of the study. The section comprises a review of the development and use of contaminated land risk assessment in European countries. The aim of the review is to identify commonly used risk assessment “systems” for comparison during later phases of this study.

A risk assessment “system” is defined as any methodology or software package designed to qualify or quantify the risk posed by a contaminant in evaluating a source-pathway-receptor linkage. A system may comprise a suite of risk based screening values, a methodology with recommended or prescribed algorithms, or a computer based package that allows site specific risk assessment to be carried out.

4.2. European Risk System Review

A literature review was undertaken for the fifteen countries currently in the European Union. In addition Norway and Switzerland were included as countries whose methodologies may significantly contribute to this study given the developmental stage of their environmental policies, highlighted by their inclusion in the CARACAS books (Ferguson & Kasamas (editors), 1999).

The reviews have been conducted using information available via internet searching, reviewing policy information detailed in guidance manuals, by direct communication with regulators or by having professionals working in the industry in different countries critique the reviews prior to their inclusion in this report. Where information could not be obtained at source the CARACAS books provide the basis for the reviews. It is acknowledged that policies could have changed in the time periods elapsed between the publication of the CARACAS books, the initiation of this project and the completion of this project.

Reviews were completed for the following countries: Austria; Belgium; Denmark; Finland; France; Germany; Greece; Ireland; Italy; Luxembourg; Netherlands; Norway; Portugal; Spain; Sweden; Switzerland; and the United Kingdom. The individual reviews are presented in Appendix B.

4.3. Proprietary Systems

There are several risk assessment systems that have been developed commercially, rather than specifically to support the approach to contaminated land of a particular country or region. Many of these are in widespread use throughout Europe (and globally), including in countries that have not yet developed their own methodology. We have included three such systems in the review stage of the comparison study:

RBCA Toolkit

The Risk Based Corrective Action (RBCA) methodology developed by the American Society for Testing and Materials (ASTM, 1995) is a human health risk assessment tool (ASTM Designation E1739-95). Groundwater Services Inc (Houston, Texas, USA) developed the methodology into a software package. The “Toolkit” software is a system that includes the theory of the RBCA methodology in a usable computer-based platform.

RISC

The Risk Integrated Software for Clean-ups (RISC) was developed by BP (Spence and BP, 2001) and is a Windows based package for the multi-media risk assessment of contaminated land. The system has three broad applications including i) estimation of human health risk from contaminated media; ii) calculation of risk based clean-up levels for various media; and iii) the performance of simple fate and transport modelling.

Risk Assessment Model (RAM)

Environmental Simulations International Ltd has developed this proprietary system. The aim of the model is to assess the potential risks to water resources in line with the P20 methodology developed in the UK.

ConSim

A commercially available software package developed by Golder Associates on behalf of the Environment Agency of England and Wales. The software is designed to assess the potential risks to water resource receptors from impacted soil and groundwater sources.

4.4. Conclusions

The development and acceptance of risk assessment as a tool for the management of contaminated land is growing, however the level of development varies considerably. Furthermore it can vary considerably between different regions of a country.

In many instances the development of a specific risk assessment system has been undertaken to directly address a legislative requirement. However in other countries where risk assessment is widely used an array of different modelling software packages could be used, provided that the models are suitable for the exposure scenario in question. In both cases the systems available are being used and it is these systems that were to be compared in later phases of this study.

The following list represents the conclusion of Phase I and includes all the systems identified and considered applicable for review in Phase II of this methodology at the time of the screening process:

System Name	Origin
CLEA	UK (England & Wales)
ConSim	UK (England & Wales)
Giuditta	Italy (Milano)
JAGG	Denmark
P20-RTW	UK (England & Wales)
RAM	UK (Commercial)
RBCA Toolkit	Commercial
Report 4639	Sweden
RISC	Commercial
Risc-Human	Netherlands
ROME	Italy
SFT 99:06	Norway
SNIFFER	UK (Scotland & Northern Ireland)
UMS	Germany
Vlier-Humaan	Belgium (Flanders)

5. PHASE II – SYSTEM SCREENING

In Phase II of the study, the selected systems are evaluated to determine the extent of their capabilities and to evaluate their applicability for use in testing with the Generic Test Site (Phase III) and Case Study Test Sites (Phase IV) and thereby fulfil the aims and objectives of the study. The following tasks were completed:

1. Systems identified in Phase I are reviewed using a pro-forma questionnaire;
2. The results of the pro-forma review are compiled into a results matrix;
3. Based on this information, each system is tested via three selection criteria;
4. A list of systems to be tested in the comparative study is finalised.

5.1. Pro-forma Screening

The pro-forma questionnaire used to evaluate each of the systems is presented in Appendix C. The questionnaire includes enquiries relating to general capability (e.g. whether it can quantify human health risk or undertake fate and transport modelling) before assessing what specific types of exposure pathways are included and the system's ability to model environmental fate and transport. System capabilities in terms of support are also evaluated, for example, does the system include databases and some comment on user friendliness is allowed.

5.2. Matrix

The completed questionnaires are presented in Appendix D and the results have been input to a matrix of risk assessment system against key risk system capabilities; presented as Table 1.

Table 1 presents a rapid means of assessing the capabilities of the model and establishes which potentially key elements of a risk assessment can be assessed with each system. The early categories include whether the system is a computer-based platform for carrying out an assessment and what are the principal aim(s) of the system. The proceeding categories allow identification of which fate and transport pathways and which exposure pathways are capable of being assessed in the system. The final categories establish the support systems behind the risk assessment systems, for example allowing comparison between those models with default chemical databases or online help support.

5.3. Criteria for System Screening

To select appropriate systems for further comparison, a set of three screening criteria were defined:

1. Can the system be used for site specific risk assessment?

By site specific it is intended that the model must be capable of assessing different site conditions where the majority of the parameters are required to be input by the user. Furthermore, by site specific, it is intended that any system primarily developed for setting initial screening values only would be removed from the assessment process.

2. Can the system quantify human health risk?

For a risk assessment system to be included in the Phase III and Phase IV testing, it must be capable of calculating either a receptor point concentration or a dose concentration for use in further risk assessment.

3. Is this the sole system developed in the country of origin?

This criteria question is intended to limit duplication of effort and ensure that the comparisons are as direct as possible. In several instances, multiple systems have been developed to fulfil identical (or at least very similar) objectives. In these instances, only one system will be included. Where two or models have been developed for similar or identical purposes, only one system has been selected. The reasons for selection vary depending on the circumstances of the system development, but are expanded upon further below.

5.4. Final System Selection

Table 2 presents a matrix of Risk Assessment System against Phase II Screening Criteria. The screening criteria have been applied broadly on the basis of the results from the Phase II Proforma Questionnaire and as a result, a list of systems for inclusion in the Phase III and Phase IV testing has been derived:

- CLEA
- JAGG
- P20-RTW
- RBCA Toolkit
- RISC
- Risc-Human
- ROME
- SFT 99:06
- UMS
- Vlier-Humaan

Table 2 includes explanatory notes regarding the selection of models where Criterion 3 applied. Further detail on the decision making process for these instances is provided below.

Note 1

At the time of issue, CLEA exists as a series of guidance manuals outlining the preferred human health risk assessment methodology for deriving “Tier 1” Soil Guideline Values (SGVs) for potentially contaminated sites. The methodology has also been developed into a software tool, also with the primary aim of developing SGVs. However, the software system is not intended for use as a site-specific risk assessment tool and subsequent research suggests that some elements of the model may change significantly in the future, for example the indoor air sub-model. For these reasons CLEA was not initially included in the study. However, as a recent development in risk assessment and given that it forms the basis for assessing contaminated soils in the UK, NICOLE / ISG requested that it be added into the study in a limited capacity (i.e. where appropriate). Therefore CLEA was included where it was felt useful comparisons could be drawn. For example, volatilisation pathways have not been considered since as described in the literature review it is probable that the sub-model currently incorporated in CLEA will be changed prior to the development of SGVs for volatile contaminants.

Note 2

ConSim and RAM have not been included as, in principle, they aim to complete the same tasks as P20-RTW and could be considered as “duplicate” systems. Each of the three systems are slightly different and ConSim and RAM have some additional features to P20-RTW, but these are not considered to add significantly to the human health risk assessment comparison study.

Note 3

Giuditta has not been included as it is very similar to ROME. ROME was selected, as it is the system developed by the national environmental regulatory body in Italy (ANPA).

Note 4

JAGG and P20-RTW have been included even though they are not intended to quantify a contaminant dose or human health risk. This is because they are site-specific risk assessment tools in that they include fate and transport models, which address potential contaminant migration and predict receptor point concentrations for comparison with acceptable levels. The results of the fate and transport models can be compared to results at a similar stage from other systems that do quantify human health risk values.

Note 5

The Swedish (Rpt 4639) and Norwegian (SFT 99:06) systems are very similar and it is understood that the Norwegian model is based largely on the Swedish model, which was developed with the main purpose of calculating generic guidelines for use in Sweden. At the time of evaluation, neither system existed as a computer based platform for risk assessment purposes. However, a decision was made to develop the SFT 99:06 system into a spreadsheet model to allow inclusion into the model comparison study. This decision was made because the guidance manual states that the algorithms in the manual can be used for site-specific human health risk assessment purposes, whereas no such comment exists in the Swedish equivalent. In addition, it was apparent that the algorithms were relatively easy to put into a spreadsheet model. It should be recognised that the Swedish methodology can also be used to develop site-specific guidelines and it should be noted that the Swedish EPA is currently developing a spreadsheet model for site-specific use. Further details on differences between the Norwegian and Swedish models are provided as a supplement to Appendix B and should enable the reader to infer similarities and differences between the Swedish model and the models considered for the GTS.

Note 6

SNIFFER has not been included since, similarly to CLEA, it was designed to assist in the UK regulation of contaminated land under Part IIA of the Environmental Protection Act 1990. Furthermore, it has recently been updated to reflect the guidance in the CLEA publications. SNIFFER did not have an associated software package to assist in its implementation at the time of the system evaluation.

Each system selected for comparison in Phase III has been described in greater detail in Appendix E.

6. PHASE III – GENERIC TEST SITE (GTS)

6.1. Introduction

The environment is inherently difficult to simplify into a conceptual site model at any site. The selection of parameters relies on the professional judgment of risk assessors and is often highly subjective. By defining a single, consistent set of input values the subjectivity of site characterisation is removed and a standard approach for the comparison of outputs is developed. Sensitivity testing of input parameters provides further conclusions with respect to the impact that potential variations in site characterisation may have on output concentrations.

The goal of the Phase III testing is to allow comparison to be made between the outputs from individual risk assessment systems using a common base point. In order to achieve this goal, the following tasks have been defined:

1. Define the Conceptual Site Model

The conceptual site model for the Generic Test Site is a series of defined source pathway receptor linkages selected to enable comparison in terms of both applicability on most real sites and what is appropriate for comparison.

2. Define the Results Comparison Methodology

The defined risk assessment systems are capable of producing a vast array of information relating to the types of assessment required. This task aims to refine the potential in order to achieve the most efficient means of comparing between the systems.

3. Define the Parameter Inputs

Standardised parameters are developed to ensure that the comparison between the system outputs is valid. In many cases the algorithms are based on similar theories, but require different parameters to achieve the same overall input. In these instances the parameters were isolated and given values to produce equivalent inputs. The final input data set ensures that potential output differences due to variations in inputs are minimised.

4. Undertake Model Simulations and Review Results

Each selected source-pathway-receptor linkage is assessed using each of the defined systems. The output is collated and compared and differences explained in terms of the variations that exist in modelling approach, hard-wired parameters and differences in algorithm interpretation.

5. Undertake Sensitivity Analysis and Review Results

The sensitivity of key input parameters is tested and the variation in the output is compared to the variation in the inputs. The results of the testing are compiled and comparisons made in the light of the findings of the original model simulations.

6.2. GTS Conceptual Site Model

The Conceptual Site Model for the GTS is a series of source-pathway-receptor linkages that provide effective and reasonable comparison between output results.

In developing the model, or linkages, we have considered typical potential contaminants, typical exposure scenarios and have also made assumptions that ensure that practical comparisons can be made, for example, assuming a permeable geological medium to ensure that contaminants are able to migrate to receptor points.

6.2.1. Sources

The characterisation of a source is often largely dependent on the means by which the release occurred in the first instance. Spill and leaks at the surface are the main cause of impact for one-off point source releases, whereas impacts derived from leaking underground pipes, storage tanks or buried waste can often go undetected and lead to larger impacts often described as continuous sources. Depending on the release mechanism, soils may be impacted from the surface (where a spill or leak may have occurred) and extend downwards to greater depth under gravity. The degree of spreading will depend on the type of contaminant, the quantity spilled and the geological conditions of the ground.

At a number of potentially contaminated sites the presence and characteristics of groundwater has a major control on the distribution of impacts in the subsurface. Groundwater can be considered a receptor in its own right (if it is part of a viable aquifer) or a means of off-site migration, or as a source. Consideration of groundwater as a source is included in the GTS.

The GTS data set considers three sources:

- **Shallow Soil**, e.g. leak from an above ground storage tank;
- **Deep Soil**, e.g. from buried waste or leaking underground storage;
- **Groundwater**, e.g. as a result of landfill leachate.

In order to assess potential risks posed by the sources, contaminants of concern must be defined and characterised. The contaminants potentially present in a source can be broadly split into two categories:

1. **Threshold contaminants (e.g. non-carcinogens)**
A threshold substance is one where there is considered to be a limit below which negative health effects are not likely to occur.
2. **Non-threshold contaminants (e.g. carcinogens)**
A non-threshold contaminant is a substance for which there is no theoretical reason why a single molecular exposure should not result in a tumour or mutation.

The characterisation of a contaminant as either a threshold or non-threshold substances has important implications on the development of dose concentrations and quantifying risk levels.

In selecting an appropriate suite of contaminants of concern, additional consideration must be given to other factors that control the behaviour of substances in the environment. For the GTS, contaminants were chosen to provide the greatest range of different characteristics including consideration of the following factors:

- Mobility – meaning the ease with which a substance dissolves and permeates the soil, is dissolved in groundwater and migrates in groundwater. The relative mobility of a substance can be determined by comparing the maximum solubility and relevant partition coefficients.
- Volatility – Meaning the ease with which the substance enters the vapour phase, which is important when assessing the risks from inhalation pathways. The relative volatility of a substance can be determined by comparing between the vapour pressures.

With consideration of the above factors in mind, and the practical requirement to include substances that could be considered common soil and groundwater contaminants on a contaminated land site, the following contaminants of concern were selected to:

1. Atrazine
2. Benzene
3. Benzo(a)pyrene
4. Cadmium
5. Trichloroethene

6.2.2. Pathways

The pathway in a source-pathway-receptor linkage is the means by which the impact characterised at the source affects the defined receptor. For the purposes of this study and the wider context of contaminated land risk assessment, there is a distinction made between direct and indirect pathways. The former is a pathway that occurs direct from source to receptor and includes such pathways as ingestion and dermal contact. The latter, indirect pathways, require some form of transport to occur from the source to the receptor point before exposure can occur, for example volatilisation and vapour transport are required before a receptor can inhale the vapours. Therefore indoor air inhalation is considered an indirect pathway.

For most indirect pathways, some form of model is required to quantify the transport of the impacts from source to receptor point or characterise the environmental pathway. These are referred to as fate and transport models. The models use the information known about the source and take account of the known environmental conditions of the site and surrounding area to predict a receptor point concentration. The receptor point concentration can be compared to an “acceptable” limit to ascertain if any risk is posed, however it can also be used as a source concentration in an exposure model to ascertain human health risk

Exposure models are those models that calculate dose concentrations in human health receptors. In this study, the calculation of the dose from a person ingesting impacted soil is an example of an exposure model. The dose calculated from someone inhaling the vapours in an indoor airspace originating from a deep soil or a groundwater source also uses an exposure model.

For direct pathways, only an “exposure model” is needed to calculate the dose concentration; because the pathway is direct, the source concentration is the receptor point concentration. For indirect pathways an “environmental model” is required to predict a receptor point concentration which is then used as the source concentration in the exposure model.

6.2.3. Receptors

The focus of the study is human health risk assessment and therefore the receptors that are included in the study are human health receptors only. Other types of receptor include ecological receptors (protected habitats or species), sensitive water resources (such as public water supply abstractions) and buildings and property.

Human health receptors comprise human beings and are often differentiated only as far as being either an adult or a child receptor. Further refinement of this can involve consideration of whether the receptor is exposed under a residential or industrial scenario, specific age groups or include consideration of target organs in the body, which are more or less affected depending on the type of contaminant of concern.

For the purposes of this study the only receptor considered are Adult exposures. This decision was made following initial comparison work with the risk assessment systems and is discussed in detail in Section 6.3.2.

6.2.4. Linkages in the GTS

It was not the intention of the study to develop a conceptual site model (CSM) that reflected a potentially real situation, this approach would have omitted several key pathways. Instead, the CSM comprises a series of individual source-pathway-receptor linkages using the three different types of source outlined in Section 6.2.1.

Figure 2 illustrates the range of receptor point concentrations, dose concentrations and quantified risk levels that have been developed through the study. Figure 3 is a schematic representation (in cross-section) of the linkages included in the GTS.

Table 4 is a matrix of risk assessment systems against the defined list of output points at which comparison will be made. From Table 4 it can be seen which systems can be used to generate the desired output data sets.

6.3. Results Comparison Methodology

The GTS conceptual model has been defined, however in order to allow efficient comparison of results between the systems, several elementary factors must be considered and reviewed to produce a consistent technical basis from which results can be generated and compared. This section provides a description of these key issues and the methodologies employed to standardise their interpretation throughout Phase III.

Output is generated at three distinct levels for comparison in this study:

- Receptor Point Concentrations;
- Dose Concentrations; and
- Quantified Risk Levels.

6.3.1. Receptor Point Concentration Comparisons

Fate and transport models are included in a number of the risk assessment systems to quantify the environmental pathway and generate Receptor Point Concentrations. Environmental pathways are simulated by a series of algorithms to enable predictions of concentrations of the contaminants at points distant from the source areas. The models fall into three types:

1. Vapour intrusion models
Receptor point concentrations have been derived in indoor air spaces from the Deep Soil and Groundwater sources into residential properties.
2. Groundwater fate and transport models.
Receptor point concentrations have been derived at hypothetical groundwater monitoring wells at a distance of 50m from the site from the Deep Soil and Groundwater sources.
3. Surface water mixing models
Receptor point concentrations have been derived for a hypothetical surface water receptor at 50m from the site that includes the effects of dilution and mixing in the surface watercourse. This fate and transport model is used in conjunction with the deep soil to groundwater model.

In total there are 5 output data sets for comparison at the *Receptor Point Concentration* level.

6.3.2. Dose Concentration Comparisons

The calculation of dose concentrations requires an input source concentration which is either directly input as in the case of direct exposure pathways (such as a soil concentration for dermal contact with soils), or is provided as an output from a fate and transport model as in the case of indirect exposure pathways (such as predicted indoor air concentrations in the inhalation of vapours in indoor air spaces). The exposure scenarios included in the GTS are as follows:

1. Direct Exposure Pathways from Shallow Soil
The direct exposure pathways include ingestion of impacted soil, dermal contact with impacted soil, ingestion of vegetables grown in impacted soil and inhalation of dust generated from impacted soil.
2. Indirect Exposure Pathways from Deep Soil
The indirect exposure pathways stem from those receptor point concentrations derived using the Deep Soil source and include inhalation of vapours in an on-site indoor air-space. And the exposure pathways associated with surface water receptors, for example dermal contact through bathing.
3. Indirect Exposure Pathways from Groundwater
The indirect exposure pathways stem from those receptor point concentrations derived using the Groundwater source and include ingestion of groundwater at a point 50m from the site.

Therefore there are 7 output data sets for comparison at the *Dose Concentration* level and these include exposure models both with and without modelled source concentrations.

Dose concentrations may be averaged over a lifetime for non-threshold compounds but averaged over a given exposure duration for threshold compounds. However, in some systems this distinction is made at the dose stage, whilst for others it is made at the risk calculation stage (Section 6.3.3). In order to ensure like with like comparisons, the dose concentrations are all averaged over the specified exposure duration. In some systems the model output was multiplied by a correction factor to account for this.

In examining the differences between the dose concentrations for child and adult receptors it became apparent that there were two key differences that would explain variations seen between the dose concentration results for the child and adult receptors. Discussing the full comparison for child receptors and adult receptors would result in repetition. Therefore the key differences are discussed below and the full comparison is conducted only for adults. The adult receptor was selected in preference to the child receptor, as it is more uniformly defined between the systems, therefore the results are subject to less differences due to hard-wired parameters.

1. Bodyweight

In the GTS, the bodyweight is defined as 20kg for a child and 70kg for an adult. These values are the hardwired defaults in UMS and were selected as the GTS values to allow direct comparison with other systems where body weight can be defined. In Vlier-Humaan, the child bodyweight is hardwired as 15kg (adult body weight is also hardwired, but at the 70kg used in the GTS dataset). Therefore it should be noted that there would be a difference in output if child dose calculations were compared between Vlier-Human and the other systems using the GTS data set.

2. Intake Rate

The intake rate also varies between child and adult receptors, for example a child may ingest more soil than an adult. The intake rate varies between exposure pathways; the selected values for each pathway are detailed in Tables 6b and c.

6.3.3. Quantified Risk Level Comparisons

Where an exposure pathway is assessed to provide a dose, it is logical to compare this dose with some form of acceptable dose to determine if a risk is presented. This is a difficult aspect to compare between the systems as the differences between the models often reflect differences in national policy on assessment of health effects.

In the first instance there is no globally agreed acceptable criteria and no globally agreed methodology of calculating a quantified risk level. In the second instance, each of the systems evaluated calculates quantified risk levels in a different way, for example some models combine pathways and others presented separately; or some systems assess dose concentrations for non-threshold substances using cancer slope factors and others use reference doses.

Therefore, the main comparisons and conclusions with respect to the Quantified Risk Levels relate to what each of the individual systems do, i.e. what are their capabilities and what are the characteristics of the outputs produced? Therefore the Quantified Risk Level comparisons focus on the following rather than specific output values:

- Output presentation
- Approach to Averaging Times
- Approach to Toxicity

6.3.4. Sensitivity Testing

To test the sensitivity of the systems to variations in parameters away from the GTS dataset, the array of input parameters have been divided into four categories:

- Physical Input Parameters

The key input variable that affect the calculation of receptor point concentrations have been modified to ascertain the variation in the results compared to the GTS results. Parameter changes only affect those models used for calculating indoor air concentrations and groundwater concentrations. The following variables have been included:

Indoor Air Concentrations:

- Crack fraction;
- Pressure Difference
- Building Height
- Ventilation Rate

Groundwater Concentrations:

- Infiltration Rate
- Groundwater Velocity
- Degradation Rate

- Chemical Input Parameters

Selected GTS results have been compared to the equivalent results where the chemical input parameters specified in the GTS have been replaced by the default chemical input parameters presented in the systems. The specific pathways that this has been carried out for are as follows:

- The receptor point concentrations for Benzene for the “Soil to Indoor Air” and “Soil to Groundwater at 50m” pathways;
- The dose concentrations for benzene and cadmium for the “dermal contact with soil” and “ingestion of vegetables” pathways.

- Toxicological Input Parameters

Selected GTS results have been compared to the equivalent results where the toxicological input parameters specified in the GTS have been replaced by the default toxicological input parameters presented in the systems. The specific pathways that this has been carried out for are as follows:

- All exposure pathways using benzene (non-threshold substance);
- All exposure pathways using TCE (threshold substance).

- Exposure Input Parameters
Selected GTS results have been compared to the equivalent results where the exposure factor input parameters specified in the GTS have been replaced by the default exposure factor input parameters presented in the systems. The dose concentrations for benzene and TCE have been tested for the following specific pathways:
 - Ingestion of Soil;
 - Inhalation of Indoor Air;
 - Ingestion of Vegetables.

6.4. Input Parameter Selection

6.4.1. Chemical and Toxicological Input Parameters

A list of the key properties for each contaminant is presented in Table 3 including source concentrations for the GTS.

A number of the risk assessment systems have chemical databases included within the software and most of the systems allow the addition or modification of new chemicals. The chemical database in the Danish JAGG model is a hardwired database linked to the model algorithms and therefore to allow comparison of results between systems, those parameters used by the JAGG model were used as the basis for the GTS chemical input dataset. As chemicals could not be added to the JAGG database, atrazine could not be added and results could not be obtained. For those parameters not required in JAGG (including the characteristics of atrazine) sources of information included those values most commonly stated in the other databases.

The toxicological dataset comprised the default inputs from the ROME software as initially it was believed that the database could not be modified, however in a later revision to the model used, modifications could be made where required.

Source Concentrations

For the soil and groundwater sources, concentrations have been defined for each contaminant of concern. The source concentrations have been set at such levels to ensure that meaningful results are achieved in the resultant concentrations for comparison. Any comparison of results between contaminants must take into consideration the differing source concentrations.

6.4.2. Physical Parameter Selection

Where fate and transport models are used, a range of parameters is required to describe the environment. These fall into the following categories:

- Source Dimensions;
- Unsaturated Zones Parameters;
- Saturated Zone Parameters;
- Building Parameters for Slab-on-grade House;
- Dust Parameters
- Receptor Point – Distance to monitoring wells;
- Surface Water – Properties of surface water receptor

A list of key physical input parameters is presented in Table 5, where additional physical input parameters are required in a single risk system they are listed in Appendix F.

In determining to values to assign to the required parameters, the following considerations were made:

- In certain pathway assessments, an input variable in a particular system was fixed or hard-wired at a specific value. In these instances the GTS data set was set to this value.
- In several instances more than one system had the same parameter hard-wired but at different values. In these instances, one of the values was used in the GTS data set, and the effect of the different hard-wired values noted in the pathway assessments
- Where parameter inputs were capable of being input independently in all models, the choice of value selected for the GTS was made on professional judgement, considering the scenario being assessed.

6.4.3. Exposure Parameter Selection

A list of key exposure input parameters is presented in Tables 6a, b and c, where additional exposure parameters are required in a single risk system the parameters are listed in Appendix F.

7. PHASE III – OUTPUT AND RESULTS

The results of the Phase III testing are provided in the form of Figures that include the tabulated results, a graphical chart and the appropriate algorithms where appropriate. The explanations of the results and the interpretations of the findings are provided in the following subsections. The text is re-produced in the Figures for completeness.

7.1. Shallow Soil Exposure Pathways

This section presents the Dose Concentration results for the following pathways:

1. Ingestion of Soil,
2. Dermal Contact with Soil,
3. the Ingestion of Vegetables (grown in shallow soil); and
4. Inhalation of Dust (derived from shallow soil).

The findings of the results comparisons are presented below.

7.1.1. Soil Ingestion

Figure 4 presents the results of the “Soil Ingestion” pathway testing. All the systems use very similar algorithms for this pathway and the results are therefore generally similar also, however there are some key differences, which are described and explained below.

CLEA predicts doses approximately four times greater than the other systems, accounted for by the default exposure parameters, which are hardwired at different values to the GTS exposure parameters. In addition, it should be noted that the CLEA model automatically includes an “indoor ingestion of dust” pathway whenever soil ingestion is selected. However this pathway only produces small doses, which are generally insignificant in comparison with the ingestion of soil dose.

Vlier-Humaan predicts doses approximately ten times smaller than the other systems. The parameter that defines the amount of soil ingested per event (AID) is hard-wired such that when it is multiplied by an exposure factor (which is an input parameter that can be varied in the model but is specified as a default number) a pre-defined average daily ingestion rate (sourced from literature) is calculated. In the GTS, AID could not be altered but the exposure factor was altered to comply with the GTS. Changing exposure without changing AID conflicts with the way in which the algorithm would typically be used.

7.1.2. Dermal Contact

Figure 4 also presents the results of the “Dermal Contact with Soil” pathway testing. The results of this pathway assessment have produced a varied response. In order to aid description the systems have been divided into three groups based on the way in which this pathway is modelled.

Group I - RISC, RBCA, ROME, SFT and UMS

These systems contain algorithms based on the fraction of contamination absorbed by the skin, defined as 0.1 in the GTS. The results are very similar across the dermal contact models, which is expected since the algorithms are essentially the same. RISC and UMS have additional factors defined, for example, a bioavailability factor in soil (RISC) and reduction factors to reduce dose if the area of ground is covered/“sealed” in hardstanding (UMS), which were set to unity in the GTS.

Group II - Risc-Human and Vlier-Humaan

A contaminant absorption rate through the skin is defined in these dermal contact models, leading to dermal doses, which are dependent on exposure time (2 hours/day in the GTS). The algorithms are similar, although the adherence factor and absorption rate are hardwired in Vlier-Humaan with values approximately 10 times more and 10 times less (than GTS values) respectively. This has the effect of producing comparable doses between Risc-Human and Vlier-Humaan.

Group III - CLEA

The dermal contact algorithm in CLEA is absorption-rate dependent, but also includes a mass-balance to account for the variability in the volatility of different compounds. The doses for the most volatile compounds considered, benzene and trichloroethene, are orders of magnitude smaller than for the other systems, reflecting the fact that these compounds volatilise at a rate that limits the residence time on the skin and therefore the amount of absorption through the skin.

7.1.3. Ingestion of Vegetables

Figure 5 presents the results of the “Ingestion of Vegetables” pathway testing.

In general, the systems calculate the ingestion of vegetable pathway through a two-step process: i) transfer of contaminants from the soil into the vegetables (root and stem) using a Bio-Concentration Factor (BCF); and ii) a proportion of the contaminated vegetables are then consumed. In the GTS dataset, the vegetables consumed were assumed to comprise 50% root vegetables and 50% stem/leaf vegetables.

The algorithms to calculate BCFs and vegetable intake are generally similar between the systems evaluated, however, several differences were encountered, which produced the variability seen in Figure 5.

- RISC generally predicts lower doses when compared with the other systems. This is due to the inclusion of an additional factor of 0.01 in the standard Briggs equation utilised by the majority of the systems to calculate the BCF for root vegetables. This factor is included to account for the differences in the structure of the barley shoots used by Briggs (Briggs *et al.* 1982) in his experiments and the structures of more typical root vegetables such as carrots and potatoes.
- UMS combines the BCFs for root and stem vegetables into a single algorithm. Within this algorithm the contribution of root and stem vegetables is hardwired as 85% stem, 15% root and there is an additional factor of 50% to account for vegetables prepared in the kitchen being washed prior to eating. The GTS dataset has stem and root consumption as 50% each. As a result of the combined algorithm and hardwired factors, the doses predicted by UMS are generally slightly lower.

- The F_{oc} value in Vlier-Humaan is slightly higher than for GTS due to the restricted input range for the system (0.0058 compared with the GTS value of 0.005). This results in a reduced BCF, of about 15% in comparison to Risc-Human.
- For metals Vlier-Humaan has an additional factor to account for the soil pH, resulting in the cadmium stem and root BCFs being divided by two in comparison to Risc-Human and lowering the plant uptake. Vlier-Humaan predicts a porewater concentration for cadmium that is slightly larger than that predicted by Risc-Human, which has the effect of increasing the plant uptake, thereby canceling some of the effects of the earlier differences. In Vlier-Humaan the ingestion rate is hard-wired at a lower value than the GTS (0.345 kg/day compared with 0.475 kg/day) reducing the vegetable ingestion dose in comparison to Risc-Human by a total factor of approximately 30-40%.
- CLEA splits vegetables into six groups, as opposed to two in the other systems. The BCFs are defined in a similar way to the other vegetable ingestion models, however the factor to calculate porewater concentration is not allowed to exceed unity. This has the effect of reducing the benzene and trichloroethene doses in comparison to other systems. The cadmium BCFs for root and stem vegetables are hard-wired in CLEA at values higher than the GTS, explaining why the dose is higher than for the other systems.
- The ingestion rate and fraction of home-grown vegetables consumed are probabilistic parameters in CLEA, making direct comparison with the remaining systems difficult. However, it is anticipated that this aspect is the explanation behind any differences that are not attributable to the variation in BCF.

7.1.4. Inhalation of Dust Outdoors

Figure 6 presents the results of the “Inhalation of Dust Outdoors” pathway testing.

In quantifying the dose received via the inhalation of dust outdoors, it is assumed that the impacted soil is superficial. A two-group distinction can be made between the approaches of the systems to the evaluation of this pathway. Group I is more simplistic and comprises all the systems capable of assessing this pathway with the exception of RBCA and CLEA, which form Group II.

Group I – Risc-Human, Rome, SFT99:06, UMS, Vlier-Humaan

The dust inhalation models in the Group I systems require the user to specify a dust concentration, i.e. the number of dust particles per cubic metre of air. A fraction of the dust particles are then specified as being contaminated and the dust is then inhaled at a rate specified in the GTS as $2\text{m}^3 / \text{hr}$.

SFT99:06 predicted higher doses than the other systems as it assumes that inhalation takes place all day (i.e. the algorithms do not include an exposure duration time), compared to the 2 hrs/day defined for the GTS.

The fraction of dust that is contaminated, defined as 0.5 in the GTS, is not a factor in the SFT99:06 or UMS algorithms. UMS includes two additional factors, one that reduces the dose by allowing for a decreased dust concentration due to precipitation (a hardwired factor of 0.66 is incorporated), and one that assumes that a hardwired percentage of 40% of the dust present is inhaled. In addition, for organic compounds with a Henry's Law Constant greater than 0.01 (atrazine and benzo(a)pyrene) the contaminant concentration is multiplied by a “soil to dust transfer factor” of 8. For inorganic contaminants (e.g. cadmium) this factor is 4.

Of the factors discussed above, some will result in an increased dose and some will result in a decreased dose, with the overall balance generally resulting in higher doses than Risc-Human, ROME and Vlier-Humaan.

The algorithms in Risc-Human, ROME and Vlier-Humaan would lead us to expect similar results for these three dust inhalation models. However, Risc-Human produces higher doses than Vlier-Humaan, but lower doses than ROME; the following explanations were identified:

- The lung retention factor of 0.75 is not included in ROME, resulting in slightly higher doses.
- The GTS inhalation rate for outdoor exposure was specified as 2.0m/hr, whereas in Risc-Human the user can only input a maximum value of 1.67m/hr.
- The GTS inhalation rate for outdoor exposure was specified as 2.0m/hr, whereas in Vlier-Humaan, the inhalation rate is hardwired as 0.83 m/hr.

Group II - RBCA and CLEA

The RBCA and CLEA algorithms are more complex being based on particulate emission and the lateral dispersion of the dust particles, however the algorithms used by RBCA and CLEA are different. Furthermore the input of GTS parameters was not possible with CLEA as all outdoor air parameters relating to particle emission and lateral dispersion are hardwired in CLEA and could not therefore be set at the same values as used in RBCA.

The risk assessment methodology employed by RBCA for inhalation pathways calculates risks levels through the comparison of the predicted “contaminant” concentration in the air space against a defined “acceptable” concentration in air as opposed to calculating the risk by comparing the calculated dose to a RfD. For the “inhalation of dust outdoors” pathway the concentration in outdoor air/dust predicted by RBCA has been converted to a dose to allow direct comparison with other system outputs. The hand calculation used incorporates the relevant exposure values from the GTS with the exception of the exposure duration, where 24hrs/day has been used as opposed to the 2hrs/day in the GTS. It should be stressed that the conversion by hand calculation was carried out solely to make the results comparable with the other system outputs and that the RBCA output would not, in practice, be used in this way.

7.2. Deep Soil Environmental and Exposure Pathways

This section presents the results and findings of the following Deep Soil “environmental” and “exposure” pathway assessments:

Soil to Indoor Air

1. Deep Soil to Indoor Air - Receptor Point Concentrations in Indoor Air;
2. Exposure to Indoor Air by an Adult Receptor - Dose Concentrations in Indoor Air.

Soil Leaching to Groundwater

3. Leaching to Groundwater, Migration to a 50m receptor well – Receptor Point Concentration in Groundwater at 50m;
4. Groundwater leakage and mixing in Surface Water – Receptor Point Concentrations in Surface Water;
5. Exposure to Surface Water by an Adult Receptor – Dose Concentration in Surface Water.

7.2.1. Receptor Point Concentration in Indoor Air

Figure 7 illustrates the predicted receptor point concentrations of the Soil to Indoor air fate and transport pathway and Figure 8 presents a simplified process flow diagram illustrating how each system undertakes the modelling procedure and where differences in system set-up occur.

The range of receptor point concentrations reported for this pathway spans over 3 orders of magnitude, reflecting the wide range of methods employed to address this pathway, as illustrated on Figure 8. Discussion of the results is presented in the following subsections.

UMS

The UMS system predicts a greater indoor air concentration in comparison to the other indoor air models. UMS assumes that the indoor air concentration is 1 % of the predicted soil gas concentration. This “transfer factor” cannot be modified.

The remaining indoor air models incorporate fate and transport models, accounting for environmental and building characteristics.

SFT99:06

The Norwegian system applies a dilution factor to the pore air concentration to calculate the indoor air concentration. The user is required to enter an intrusion rate of pore air into the building. This intrusion rate was taken to be the same as the intrusion rate calculated within the RISC software. Therefore, the SFT99:06 results are similar to the RISC results.

JAGG

The JAGG system calculates air transport through cracks according to Baker, Sharples & Ward (reference, page 207, JAGG Manual). Flow through a concrete deck is calculated by means of a “Cubic law” (equation 44, page 201, JAGG Manual). This algorithm is significantly different to the Johnson and Ettinger algorithm (see below – RISC & RBCA) and appears to give a more conservative prediction for volatile compounds whilst still giving results of the same order of magnitude. The JAGG system also includes a complex method for calculating the crack fraction, however the calculated crack fraction in JAGG was manipulated to equal the crack fraction for other systems.

RISC & RBCA (Group I)

Both of these systems include indoor air models based on the Johnson and Ettinger algorithm for predicting vapour intrusion rates. However, RBCA incorporates two alternative expressions for the volatilisation factor from subsurface soils to enclosed spaces, (VF_{se}, equations CM-4a and CM-4b, page B-3, RBCA Toolkit Manual) whereas only the first algorithm is incorporated in RISC (equation D-6, page D-6, RISC Manual). The first algorithm (CM-4a) models an infinite source and the second models a finite source assuming that volatilisation occurs at a constant rate over the defined exposure period. RBCA automatically selects the expression that results in the lowest VF_{se}, which means for the more volatile compounds (i.e. benzene and TCE) the finite source algorithm is adopted by RBCA, leading to a difference between RBCA and RISC which uses the infinite source. For less volatile compounds, i.e. atrazine and benzo(a)pyrene, RBCA defaults to the infinite source algorithm, giving RBCA similar predicted indoor air concentrations to RISC.

Risc-Human and Vlier-Humaan (Group II)

The algorithms used to calculate soil air concentration and flux to the house foundation are similar in Risc-Human and Vlier-Humaan. However there are a number of differences between the systems that require explanation.

- When modelling the indoor air pathway Risc-Human automatically calculates an indoor air concentration and an outdoor air concentration and adopts the maximum as the receptor point concentration. In the results of the GTS, Risc-Human adopts the predicted outdoor air concentration for atrazine and benzo(a)pyrene, and as a consequence these concentrations are not directly comparable between Risc-Human and any of the other systems, including Vlier-Humaan which uses the predicted indoor air concentration for all compounds.
- The indoor air pathway defined for the GTS models a “slab on grade” style construction, however Risc-Human is not designed to model a house without a basement or crawl space. Therefore to take account of this, the input parameters have been modified for Risc-Human to generate a similar scenario as the GTS. A basement was modelled with the same dimensions as the indoor air space, with the soil source the same depth below the basement foundation as in the GTS. Risc-Human also includes two air-mixing algorithms to calculate the indoor air concentration from a soil source. The algorithms differ to take account of the relationship between the basement depth and depth to soil source from the surface. However, for the GTS scenario Risc-Human and Vlier-Humaan appear to use the same algorithm.
- In Vlier-Humaan the range of F_{oc} values that can be entered for the partitioning calculation is 1-100% organic matter content, compared with the GTS value of 0.86. This difference results in a decrease in the predicted indoor air concentration.
- In Vlier-Humaan the depth to the soil source from the surface is hard-wired at 0.75m, compared with the GTS value of 1.0m. This difference results in an increase in the predicted indoor air concentration.

Group I v's Group II

The major difference between the Group I and Group II systems is that Group I systems predict the indoor air concentration by modelling an intrusion rate through cracks in a concrete foundation for both diffusive and advective mechanisms (based on the Johnson and Ettinger algorithms) and Group II systems model an intrusion rate through pores in a concrete foundation for diffusion only (based on CSOIL algorithms).

7.2.2. Dose Concentration from Inhalation of Indoor Air

Figure 9 presents the dose concentrations calculated via the inhalation of indoor air. The indoor air concentrations were derived from the receptor point concentrations derived in Section 7.2.1 (Figure 7).

The majority of the differences between the doses are attributable to the different indoor air concentrations being used in the receptor point calculations (Section 7.3.1). For example, the predicted dose from UMS is generally at least one order of magnitude greater than that predicted by the other systems, and the indoor air concentrations predicted by UMS differ by a similar order of magnitude.

The risk assessment methodology employed by RBCA for inhalation pathways calculates risks levels through the comparison of the predicted “contaminant” concentration in the air space against a defined “acceptable” concentration in air as opposed to calculating the risk by comparing the calculated dose to a RfD. For the “inhalation of indoor air” pathway the concentration in indoor air predicted by RBCA has been converted to a dose to allow direct comparison with other system outputs. The hand calculation used incorporates the relevant exposure values from the GTS with the exception of the exposure duration, where 24hrs/day has been used as opposed to the 16hrs/day in the GTS. It should be stressed that the conversion by hand calculation was carried out solely to make the results comparable with the other system outputs and that the RBCA output would not, in practice, be used in this way.

Two differences in the dose algorithms have been identified across the systems, which relate to differences in the exposure modelling:

- RISC is the only system to include a lung retention factor of 0.75 for the inhalation of indoor air. This factor is incorporated into other systems for the pathway of inhalation of dust.
- SFT models inhalation of indoor air over a 24 hour period, as opposed to the 16 hours/day in the GTS exposure scenario. Therefore the SFT doses are generally greater than for other comparable systems.

7.2.3. Groundwater Concentration at 50m from Deep Soil Source

This comparison examines the maximum predicted groundwater concentrations at 50m as a result of leaching (via infiltration) of the soil source and subsequent groundwater migration. The results of the modelling are presented as Figure 10. Simplified schematics of the models are presented as Figure 11. The reader is also directed to Section 7.3.2 where the results of the “Groundwater Source to the Groundwater Receptor” at 50m are presented.

The algorithm used to describe partitioning between soil and soil porewater is the same in all systems. Differences in receptor point concentrations are derived from the mixing of soil porewater in the underlying groundwater followed by attenuation during subsequent groundwater migration to the 50m point.

Group I - RBCA, P20 and ROME

The predicted receptor point concentrations from these systems are in general agreement. In each case, both the algorithm used to dilute soil porewater in groundwater and the algorithm used to model groundwater transport to a receptor point (which includes attenuation and dispersion in three dimensions) are similar.

RISC

The predicted receptor point concentrations from RISC are lower than P20, RBCA and ROME. RISC is similar to the other three systems in many ways, for example, the partitioning equation from soil to soil porewater is identical to the Group I systems. Furthermore the groundwater fate and transport model simulates one-dimensional advection and three dimensional dispersion, which is also identical to the Group I systems. However there are also a number of significant differences between RISC and the Group I systems:

- The solution of the three-dimensional dispersion equation for groundwater fate and transport in RISC is a semi analytical solution being numerical in time and analytical in space, and is therefore different from the other models.

- The maximum duration of groundwater flow is specified as 100 years in RISC as opposed to infinity for the Group I systems. 100 years is an insufficient time for benzo(a)pyrene to migrate to the receptor point.
- The characterisation of the Receptor Point is more complex in the RISC system. The details of the receptor well are included with respect to the position of the well screen and the number of averaging points to be used in calculating the concentration. The default values were used in the GTS (only used by RISC) and have the effect of reducing the predicted concentration at the receptor point.
- The saturated soil source model in RISC is a time dependant finite source compared to the infinite sources used in Group I. This has the effect of limiting the quantity released over time as a source concentration for the groundwater fate and transport model and therefore limits the predicted receptor point concentration. It is anticipated that the overall effect of reducing the concentrations may be symptomatic of the nature of the GTS conceptual model and that under different conceptual conditions the relationship between the systems could be different.
- The mixing zone depth is calculated differently in RISC resulting in the use of a smaller mixing zone depth. This has the effect of reducing predicted receptor point concentrations by RISC, however, in reducing the mixing zone depth in P20, the dilution factor is reduced and the predicted concentrations increases. It is anticipated that the way in which the source term is handled in the Soil to Groundwater model affects the sensitivity of the remaining input parameters.

SFT99:06

SFT also predicts lower concentrations than the Group I systems. The receptor point groundwater concentration is calculated via a “soil porewater to groundwater dilution factor” that incorporates the distance to the receptor point. The mixing zone depth, d_{mix} , is dependent on the distance to the receptor point and increases with distance. The lateral flow increases due to infiltration between the site and the compliance point. The mixing zone depth calculation is identical to that used in P20 and ROME when the distance to the receptor point is removed from the equation. No further attenuation within groundwater is included.

JAGG

The algorithm used in JAGG is conceptually similar to that used in SFT99:06, however it does differ in the treatment of the mixing zone depth (page 231, JAGG manual). JAGG also does not allow the user to input the distance to the receptor point. The concentrations of the contaminants are usually calculated for a point at a distance from the pollution source that corresponds to one year of groundwater flow, limited to a maximum of 100m. However, interim results are presented in the model for a range of distances, including 50m. Therefore it has been possible to obtain results for the comparison. In addition, it is not possible to specify the longitudinal dispersivity, which may therefore differ from the 2m specified for the generic test site. Standard values given in Appendix 5-8 of the JAGG manual appear to estimate a longitudinal dispersivity of 0.008m for a 50m distance.

7.2.4. Surface Water Concentrations

Figure 12 presents the predicted surface water concentrations in a hypothetical receiving watercourse located 50m from the impacted soil source. The source concentrations for these models comprise the output receptor point concentrations from the Deep Soil to Groundwater model (Section 7.2.3). The impacted groundwater at 50m from the site is assumed to discharge into the watercourse.

Three of the systems selected for Phase III testing are capable of generating surface water concentrations. Figure 13 presents simplified process flow diagrams for how each risk assessment system undertakes the modelling of this pathway. The inclusion of this pathway was under development in the ROME software at the time of evaluation and it is anticipated that the pathway will be included in later versions of the model. Risc-Human and Vlier-Humaan include surface water exposure models, however both systems model a surface water feature located adjacent to the site and are based on impacted sediment washing into the surface water feature as the mechanism by which impact occurs. They do not include groundwater migration and therefore do not comply with the GTS scenario in question.

It should be noted that the variations in receptor point concentrations identified at the surface water will be heavily influenced by the variations in the concentrations in the groundwater at 50m from the site.

RBCA and RISC

The source concentrations for the surface water mixing models were lower for RISC than for RBCA, however the predicted surface water concentrations are in reasonable agreement. The same algorithm is used to calculate the mixing of groundwater in surface water in both systems, however the parameters within this algorithm are defined slightly differently. This results in the predicted *surface water* concentrations in RISC and RBCA being much closer than the predicted *groundwater* concentrations at 50m (Section 7.2.3 and Figure 10). The width of the source at the surface water is calculated in RISC following the equation (Domenico, 1987: Page L-12, RISC Manual) to estimate the distance off-the-centreline at which the concentration becomes 5% of the centreline concentration. In RBCA, however, the user must define the width of source as a parameter value. In the GTS the width was estimated based on the on-site source width and the lateral dispersion coefficient. This results in a slightly higher width in RISC, which results in less mixing and consequently a higher surface water concentration.

SFT99:06

SFT predicts a lower concentration in the groundwater at 50m than the other systems, and correspondingly predicts a lower surface water concentration. However, the relative differences between the model outputs in predicted surface water concentrations are not as pronounced as for the predicted groundwater concentrations. The SFT mixing algorithm calculates a larger mixing zone depth resulting in a smaller groundwater to surface water dilution factor, and therefore a higher predicted surface water concentration.

7.2.5. Surface Water Exposure - Dose Concentrations

Figure 14 presents the doses calculated as a result of exposure pathways associated with the impacted surface watercourse outlined in Section 7.2.4.

RISC and RBCA are able to model the ingestion and dermal contact of surface water and RBCA and SFT99:06 can model ingestion of fish caught from the impacted surface water.

Risc-Human and Vlier-Humaan include surface water exposure models, however both systems model a surface water feature located adjacent to the site and are based on impacted sediment washing into the surface water feature as the mechanism by which impact occurs. This does not comply with the GTS scenario in question and the models have been omitted from this pathway assessment.

Ingestion and Dermal Contact

The predicted dose concentrations for RISC and RBCA are comparable for atrazine, benzene and TCE. The small difference between the doses is largely attributable to the differences in surface water receptor point concentrations as the algorithms for calculating the ingestion and dermal contact with surface water doses are similar.

Ingestion of Fish

The surface water concentrations predicted by RBCA are approximately double those predicted by SFT (discussed in Section 7.2.4). However for ingestion of fish the dose concentrations predicted by SFT99:06 are significantly larger than those calculated by RBCA. This is anticipated to be a region specific difference in that SFT99:06 includes an “exposure to fish” factor of 350 days per year. RBCA requires the input parameter of fish ingestion in kg/day, however no exposure frequency for ingestion of fish is defined. This results in the dose concentration of RBCA being approximately 150 times lower than SFT99:06 (the difference would be a factor of 350 if the starting concentrations in surface water were equal). The dose algorithms are otherwise the same in RBCA and SFT.

7.3. Groundwater Environmental and Exposure Pathways

This section presents the results and findings of the following groundwater source “environmental” and “exposure” pathways:

Groundwater to Indoor Air

1. Groundwater to Indoor Air - Receptor Point Concentrations in Indoor Air;

Groundwater to Groundwater

2. Groundwater source calculating Receptor Point Concentration in Groundwater at 50m;
3. Ingestion of impacted groundwater at 50m.

7.3.1. Receptor Point Concentration in Indoor Air

Figure 15 presents the predicted indoor air concentrations in a hypothetical indoor airspace that is situated directly above the impacted groundwater source. Figure 16 presents simplified process flow diagrams of how each risk assessment system undertakes the modelling pathway.

JAGG was not included in this part of the comparison study as Arcadis were unable to ascertain how / if the model (only available in Danish) predicted the concentration in indoor air from a groundwater source. Following feedback from the Danish EPA subsequent to the presentation of results at the Consoil Conference 2003, Arcadis acknowledge that JAGG can be used to model this pathway.

The findings of the assessment of this pathway mirror the findings of the “Deep Soil to Indoor Air” pathway with the exception that the “source” is developed from a partitioning equation between water and air. The reader is directed to Section 7.2.1 for an explanation of the differences between the outputs from RISC, RBCA, Risc-Human and Vlier-humaan, the systems for which this pathway could be compared.

The only finding to report from this pathway is the difference in the handling of vadose advection between RISC and RBCA. As in the deep soil source model both systems include indoor air models based on the Johnson and Ettinger model for predicting vapour intrusion rates. However RISC excludes advection into the building because it is assumed that the capillary fringe diffusion resistance dominates the transport of vapours. RBCA includes the advection mechanism (i.e. that a pressure difference exists between the building and the soil gas below the building or $P > 0$). This results in a smaller predicted indoor air concentration in RISC compared with RBCA, which is more pronounced for the less volatile compounds, such as benzo(a)pyrene.

7.3.2. Groundwater Concentrations at 50m Receptor Point

Figure 17 illustrates the predicted receptor point concentrations of the groundwater migration pathway. Figure 18 presents simplified process flow diagrams of the mechanism by which each risk assessment system undertakes the modelling of this pathway.

Many of the explanations developed for the assessment of the “Deep Soil to Groundwater at 50m” pathway are applicable to this pathway assessment also (Section 7.2.3). It should be noted that the results here represent a review of the performance of the groundwater fate and transport model without interference from soil source algorithms. The results identified here can be viewed as contributing factors to the results identified in Section 7.2.3.

- RBCA, ROME and P20 predict very similar concentrations, as would be expected since all three groundwater models are based on the steady-state Domenico equation, an analytical solution of the advection-dispersion equation.
- RISC predicts slightly lower concentrations for two main reasons: 1) the well screen defaults were used to define the receptor point characteristics, this accounts of over 50% of the difference between the RISC and the models described above; and 2) RISC uses the AT123D code (Yeh, 1981), which is “semi-analytical” because it is analytical in the spatial domain but numerical in time.
- The JAGG groundwater fate and transport model uses a different approach whereby the input concentration is multiplied by a factor based on dilution in groundwater between the site and the receptor point. The results appear comparable for the selected contaminants under the conditions defined for the GTS.

7.3.3. Groundwater Ingestion Dose Concentrations

Figure 19 presents the predicted doses via the ingestion of groundwater at the hypothetical 50m receptor well. Only the RISC and RBCA systems are capable of calculating doses for exposure to impacted groundwater distant from the site.

The differences identified are largely attributable to the differences identified in the predicted groundwater concentrations (Section 7.3.2). However, it is noteworthy that the RISC groundwater model is a transient model, predicting receptor point concentrations on a year-by-year basis. The source concentration used in the algorithm for calculating the dose are not the maximum concentration but the maximum averaged concentrations over the defined exposure duration. This would result in a slightly lower concentration than the predicted maximum. However, this did not having a significant effect on the results generated for the GTS conceptual model.

7.4. Phase III Risk Calculations

7.4.1. Results Presentation

The different ways in which the systems report the risks makes it difficult to draw comparisons between the actual risk levels:

- RBCA, ROME and RISC derive a separate risk level for each of the individual pathways included for assessment.
- Risc-Human and Vlier-Humaan each report one risk level, based on all of the active pathways included for assessment.
- UMS reports three risk levels, one for ingestion, one for dermal contact and one for inhalation.
- CLEA outputs one risk level for each model run; the user must select one intake pathway for the health criteria on which the risk level is based and separate models must be run if different health criteria are defined for each intake pathway.

However, the algorithms to convert doses into risk levels are all similar, therefore any major differences between risk levels are likely to be due to differences in the dose concentrations. Furthermore, the risk levels are highly dependent on the definition of toxicity health criteria (RfDs, slope factors etc), and the means in which averaging times are used to make the health criteria applicable to the dose. The difference in the way in which the models handles these aspects is discussed below.

7.4.2. Toxicity

There are a number of approaches employed by the systems to calculate risk levels. The approaches are discussed below, using the most common nomenclature for each approach.

Carcinogenic Slope Factor (CSF) – for non-threshold contaminants

The predicted lifetime averaged daily dose is multiplied by a CSF ($1/(\text{mg}_{\text{contaminant}}/\text{kg}_{\text{body weight}}\cdot\text{day})$) to calculate the individual excess lifetime cancer risk (IELCR).

Reference Dose (RfD) – for threshold and non-threshold contaminants

The predicted daily dose is divided by an RfD ($\text{mg}_{\text{contaminant}}/\text{kg}_{\text{body weight}}\cdot\text{day}$) to calculate a risk index (RI). This calculation of risk levels is predominantly used for *threshold* contaminants. However for *non-threshold* contaminants in Risc-Human and Vlier-Humaan, an RfD is presented instead of a CSF, based on a 1×10^{-4} (Risc-Human) and 1×10^{-5} (Vlier-Humaan) lifetime excess cancer risk. The predicted lifetime averaged daily dose is divided by the RfD to calculate an RI, as for threshold contaminants. In CLEA, *non-threshold* contaminants are also assessed using an RfD. However, the definition of an acceptable level of risk differs slightly from Risc-Human and Vlier-Humaan as the RfD is compared to the predicted daily dose, as opposed to the predicted lifetime averaged daily dose.

Unit Risk Factor (URF) – for non-threshold contaminants

For inhalation exposure pathways within RBCA, a dose is not calculated. Instead, the receptor point concentration – the predicted air concentration– is multiplied by a URF ($1/(\text{mg}_{\text{contaminant}}/\text{m}^3_{\text{air}})$), providing an IELCR as for the slope factor approach.

Tolerable Concentration in Air (TCA) – for threshold contaminants

As for inhalation exposure to a non-threshold contaminant, a dose is not calculated for inhalation in RBCA. Instead, the receptor point concentration – the predicted air concentration – is divided by a TCA ($\text{mg}_{\text{contaminant}}/\text{m}^3_{\text{air}}$), providing an RI as for the reference dose approach. Risc-Human and Vlier-Humaan both provide an RI based on a TCA, as well as an RI based on an RfD, for inhalation exposure pathways, using the methodology for assessing non-threshold and threshold contaminants described above.

The table below summarises the toxicological approach for the systems assessed which are designed to produce risk levels as an output:

Toxicological Approach	System	
	Non-Threshold Contaminants	Threshold Contaminants
RfD	CLEA Risc-Human Vlier-Humaan UMS	CLEA RBCA (direct contact) RISC Risc-Human ROME UMS Vlier-Humaan
CSF	RBCA (direct contact) RISC ROME	
TCA	Risc-Human (inhalation-optional) Vlier-Humaan (inhalation-optional)	RBCA (inhalation) Risc-Human (optional) Vlier-Humaan (optional)
URF	RBCA (inhalation)	

In order to maintain consistent data inputs between the non-threshold forms of assessment, an acceptable risk level of 1×10^{-5} was chosen to allow conversion of the adopted CSF to an equivalent RfD where required. The adopted CSF for inhalation exposure was converted to a URF for RBCA using the GTS adult inhalation rate and body weight.

7.4.3. Averaging Times Risk

Quantified Risk Levels have been derived for both the child and adult receptors. For threshold substances, the exposure duration and averaging time are generally defined to be equal and therefore “fall-out” of the algorithms, such that differences in the definition of these parameters for child and adult receptors would have no effect. However, for non-threshold substances there is a variation in the averaging time, which is effected in the models in different ways:

- In RISC, RBCA and ROME, the systems include the exposure duration and develop a lifetime averaged dose for non-threshold contaminants. RISC and ROME calculate doses for a child, an adult and a combined dose.
- In SFT and CLEA the averaging time is equal to the exposure duration, i.e. there is no difference in the assessment of threshold and non-threshold contaminants, therefore *this* difference between child and adult doses would not have an effect.
- In UMS, non-threshold doses are developed using a lifetime receptor. The lifetime receptor doses are derived by weighting the child and adult doses over their respective exposure durations (5 years and 30 years respectively).

- In Risc-Human and Vlier-Humaan, the child and adult doses are calculated with the averaging time equal to the exposure duration. However, a third dose, the lifetime dose, is calculated for non-threshold contaminants, which weights the child and adult receptor doses over their respective “lifetimes”, 5 years for a child and 65 years for an adult in Risc-Human (to best approximate the GTS) and 6 years for a child and 64 years for an adult in Vlier-Humaan (hard-wired). Risc-Human varies from the other systems in that it also develops a lifetime averaged dose for threshold substances.

7.4.4. System Testing

A scenario has been defined to look at the risk level from five pathways, for each system. However, the issues discussed above are often specified by the nation in which the software is being used, therefore it is not considered appropriate to draw any conclusions on differences in risk levels that are based on an “arbitrary” standardised set of health criteria.

Risk levels have been calculated for one threshold contaminant (TCE) and one non-threshold contaminant (benzene) to demonstrate similarities and differences in the way in which each system assesses these two groups of contaminants.

Threshold Contaminants

The risk level outputs are presented in Tables 7a and 7b and the main differences in the form of the output is discussed below.

The majority of the systems calculate risk levels separately for each receptor, defined as child and adult for this scenario, with the following exceptions:

- In RBCA although the user can specify these two receptors, the output risk level is always based on an adult receptor.
- The output from Risc-Human is based on a lifelong dose (input here as the adult dose)
- The output from Vlier-Humaan automatically selects the larger of the child or adult dose, on which it bases the risk calculation; in the GTS case it is the child dose.
- In CLEA separate models need to be run for each age group, therefore the risk level could be defined separately for child and adult receptors by running two separate models, one to model a child and one to model an adult. However, the CLEA methodology states that the worst case should always be assumed, therefore the risk level is always quoted for a female, aged 0 to 6 years.

The RfD is specified separately for the three intake pathways – ingestion, dermal contact and inhalation – for all of the systems, with the exception of Vlier-Humaan where only two RfDs are defined (for oral and inhalation) and Risc-Human where only one RfD is defined for all three pathways. For TCE, the RfD for was defined to be the same for each pathway, therefore these differences will not have an effect on the risk level results.

Non-threshold contaminants

The risk level outputs are presented in Table 8 and the main differences in the form of the output is discussed below. It should be noted that where different toxicological approaches are adopted, results could not be compared.

For non-threshold contaminants, the risk levels are based on doses averaged over a lifetime, with the exception of RBCA where the risk level is based on the dose for an adult, and CLEA where the risk levels are based on a 0-6 year old female, the worst case scenario across all the age groups.

Risc-Human and Vlier-Humaan both calculate a dose for child and adult receptors separately and then weight these doses over the respective lifetimes of an adult and a child, where the sum equals the “lifetime” exposure duration. These lifetimes are defined as follows:

- Risc-Human: 5 years for a child and 65 years for an adult. (selected to best match GTS)
- Vlier-Humaan: 6 years for a child and 64 years for an adult. (hard-wired)

UMS and ROME use a similar algorithm, however the doses are weighted over the exposure durations defined as 5 years for a child and 30 years for an adult.

RISC calculates separate risk levels for the child and adult receptors and then combines them, the result of which is similar to weighting the child and adult doses over their respective exposure durations.

8. PHASE III SENSITIVITY TESTING

Sensitivity testing has been carried out on selected parameters to develop responses to changes in physical, chemical, toxicological and exposure factor variations.

8.1. Physical Input Parameters – Indoor Air

There are a number of input parameters that if varied slightly can have profound impacts upon the results of a number of the models. The parameters tested for indoor air sensitivity include Crack Fraction (and air filled porosity), Pressure Difference, Building Height and Ventilation Rate. The input parameters used in the sensitivity testing of the systems are presented in Table 9 and each parameter is discussed below.

8.1.1. Crack Fraction and Air Filled Porosity

The variation in the type of input parameters required to modify the effects on the Vapour Intrusion models has meant that a consistent approach to testing sensitivity could not be developed. Figure 20 shows the results of the testing carried out and a discussion of the findings is provided in the following sections.

RISC & RBCA

Both of these models simulate vapour intrusion via cracks. The variation in crack fraction has been investigated for both benzene and benzo(a)pyrene to address effects on volatile organic compounds (VOCs) and semi-volatile organic compounds (SVOCs). In RBCA toolkit's approach, VOCs are assessed using a different model to the less volatile SVOCs. For VOCs changes in crack fraction have no effect. For SVOCs the effects of changing the crack fraction is similar to that in RISC. For benzo(a)pyrene RISC is slightly more sensitive to the crack fraction than RBCA, however, the crack fraction appears to be a relatively insensitive parameter for both models with a 50% change leading to a maximum change of 10% in the indoor air concentration.

Risc-Human & Vlier-Humaan

Both of these models simulate intrusion via the porosity of the concrete foundation. The concrete porosity is hardwired in Vlier-Humaan so its sensitivity could not be tested. In Risc-Human, the porosity is specified by defining a solid phase fraction and an air phase fraction (thereby defining the resultant water filled fraction). When the fraction of solid phase was kept constant and the air filled fraction was varied by 50%, the indoor air concentration changed by a maximum of 119%, with an increase in air-filled porosity giving an increase in indoor air concentration. The air filled porosity in Risc-Human is a more sensitive parameter than the crack fraction in either RISC or RBCA.

SFT

The sensitivity of a crack fraction or air filled porosity was not tested in SFT as it required the definition of an intrusion rate, rather than individual parameters used to calculate the intrusion rate. Crack fraction is therefore not a required input.

JAGG

The sensitivity has not been tested in JAGG as this system calculates a crack fraction based on the weathering of the concrete used to make the foundation, as such the crack fraction cannot be simply varied.

8.1.2. Pressure Difference

The variation in pressure difference has been investigated for both benzene and benzo(a)pyrene to address the relative effects on a VOC and an SVOC. Figure 21 shows the results of the testing carried out and a discussion of the findings is provided in the following sections:

- In RBCA toolkit's approach, VOCs are assessed using a different model to the less volatile SVOCs. For VOCs, changes in pressure difference have no effect, however for SVOCs the effects of changing the pressure difference is similar to that in RISC.
- JAGG is less sensitive to changes in pressure difference than RISC.
- Where an effect occurs the three systems are more sensitive to a reduction in pressure difference than an increase in pressure difference.

8.1.3. Building Height

The effect of varying the building height has been investigated for benzene for six systems. The building height was increased and decreased by 25% (range from 1.5m to 2.5m). Figure 22 shows the results of the testing carried out and a discussion of the findings is provided in the following sections:

- RISC, JAGG, RBCA and SFT99:06 all respond in a similar way, being more sensitive to a decrease in height than an increase in height.
- Risc-Human follows this trend but is generally less sensitive.
- For an increase of 25%, Vlier-Humaan displays similar sensitivity to RISC, JAGG, RBCA and SFT99:06, however, there is no value for a decrease in building height of 25% as the minimum height that can be specified in Vlier-Humaan is 2m.

8.1.4. Ventilation Rate

The effect of varying the building ventilation rate has been investigated for benzene for six systems. The building ventilation rate was increased and decreased by 25% (ranging from 15 to 9 exchanges per day). Figure 22 shows the results of the testing carried out and a discussion of the findings is provided in the following sections:

- All of the models respond in a similar way.
- The models are more sensitive to a decrease of 25% than for an increase.

8.2. Physical Input Parameters – Groundwater Migration

Three parameters have been investigated for the Soil to Groundwater Migration model; these are Infiltration Rate, Groundwater Velocity and Degradation Rate. The input parameters used in the sensitivity testing of the systems are presented in Table 9.

Figure 23 presents the results of the testing and each parameter is discussed below.

8.2.1. Infiltration Rate

The infiltration rate is the least sensitive of the three parameters tested, with an increase in value of 20% leading to a maximum change of 12% (SFT) and a minimum change of 1% (RISC). The infiltration rate is a key factor in determining the amount of dilution that occurs when the soil porewater enters the groundwater beneath the source.

8.2.2. Groundwater Flow Velocity

The groundwater flow velocity was increased and decreased by 90% by varying the hydraulic gradient, hydraulic conductivity and effective porosity simultaneously. The main finding is that the RISC systems responds markedly differently to the other models: a 90% decrease in the groundwater velocity results in a 62% decrease in the receptor point concentration, whereas for the other systems, a 90% decrease in groundwater velocity results in an increase in receptor point concentration.

The main difference between RISC and the other models is that it is a partially transient model and as such the saturated soil source is finite and depleted in proportion to the groundwater flow beneath the site. Under the conditions set for the GTS model the increase in groundwater velocity results in a greater “source” concentration being developed in the groundwater immediately beneath the soil source and hence a greater receptor point concentration is produced. For the remaining systems, the greater the groundwater velocity the larger the dilution factor, the lower the “source” concentration in the groundwater immediately beneath the soil source and hence the lower the groundwater concentration at the receptor well.

It is likely that this variation is due to the nature of the conceptual model developed for the GTS and the relative effects of dilution in groundwater where groundwater velocities are reasonably high.

8.2.3. Degradation Rate

The different effects of inputting a nil, high-, and low-degradation rate were tested using the GTS data set.

All of the models are sensitive to degradation rate. When a high degradation rate is assumed the concentrations reduce to near zero and are therefore not seen on the graph. P20 is the least sensitive to changes in degradation rate.

8.3. Chemical Input Parameters

The sensitivity testing of Chemical Input Parameters is focused on the variations produced when the system defaults are used. The chemical properties of a substance can have an impact on the fate and transport of a substance in the environment and they can influence the amount of uptake in certain exposure pathway scenarios. For this sensitivity testing the effects of default chemical properties on the dose concentrations via dermal contact and vegetable ingestion pathways, and the receptor point concentrations in indoor air and in groundwater are tested.

The chemical property values that have been tested are presented in Table 10 and a discussion of the findings of the testing is presented below.

8.3.1. Dermal Contact with Soil

The Figure 24 presents the percentage change in dermal contact doses for benzene and cadmium. The findings are summarised below:

- RISC, Risc-Human and Vlier-Humaan show little variation from the GTS results, as parameters for this pathway in their default chemical databases are very similar to the GTS dataset.

- RBCA and UMS have higher default absorption factors for dermal uptake of benzene, explaining the increase in dermal contact doses.
- SFT and UMS have significantly higher default absorption factors for dermal uptake for cadmium, explaining the large increase in dermal contact doses.

8.3.2. Ingestion of Vegetables

The Figure 24 presents the percentage change in vegetable ingestion doses for benzene and cadmium from the GTS doses for the same substances. The findings are summarised below:

- The most pronounced change identified is the 200% increase in the dose for cadmium calculated by Risc-Human. This change has occurred due to the use of the default Bio-Concentration Factors (BCFs) as opposed to the calculation method used in the GTS testing.
- UMS and Vlier-Humaan show an increase in dose for benzene, explained by a lower default K_{oc} .
- The RISC default dataset shows little variation from the GTS dataset for benzene, with the vegetable uptake factor calculated from the chemical database. However, an increase from the GTS dose is seen in RISC for cadmium. The uptake of cadmium is defined by an input BCF value in the default database, lower than the value in the GTS dataset, explaining the decrease from GTS dose observed.
- SFT is the only system with similar chemical properties to the GTS dataset and consequently shows little variation for benzene and cadmium doses.

8.3.3. Indoor Air

Figure 25 presents the percentage change in receptor point concentrations for benzene in indoor air. The findings are summarised below:

- RBCA, RISC and ROME show little variation, as the default databases are essentially similar to the GTS dataset.
- JAGG increases by 100% due to a decrease in K_{oc} value.
- The four remaining systems show a reduced indoor air concentration, with the range of changes being from 35% for Vlier-Humaan to a 100% decrease for SFT. Which of the input parameters was most responsible for the change could not be easily determined, however the default diffusion coefficient in air is the only input parameter that is obviously different ($0.7 \text{ cm}^2/\text{s}$ compared to $0.093 \text{ cm}^2/\text{s}$) for the SFT model. The other models that predict a lower concentration also have greater diffusion coefficient in air values than the GTS dataset, with the exception of UMS, which does not use such a coefficient, but has smaller values specified for Henry's Law and K_{oc} .

8.3.4. Groundwater (from Soil Leachate Source)

Figure 25 presents the percentage change in receptor point concentrations for benzene in groundwater 50m from the source. The findings are summarised below:

- RISC, RBCA, ROME and SFT show little variation because the databases are similar to the GTS dataset.
- JAGG increases by 100%. This is likely to be due to a decrease in K_{oc} value.

8.4. Exposure Input Parameters

The sensitivity testing of Exposure Factor inputs is focused on the variations produced in Dose Concentrations when the system defaults are used. The percentage change in benzene and TCE dose concentration results from the GTS testing are provided for the Ingestion of Soil, Inhalation of Indoor Air and the Ingestion of Vegetables pathways.

The default exposure input parameter values for each system are presented in Table 11. For the RISC system, two sets of default data exist (Typical and Reasonable Maximum Exposure (RME)). In assessing the sensitivity of default exposure data the Typical Exposure data set was selected and used in the testing, though it should be noted that the RME data set more closely approximates the GTS data set.

8.4.1. Ingestion of Soil

Figure 26 presents the percentage change in Dose Concentrations for benzene and TCE for the Ingestion of Soil exposure pathway. The findings are summarised below:

- RISC shows a decrease in ingestion of soil dose for both benzene and TCE. This is explained by both a decrease in exposure duration (affecting benzene) and a decrease in ingestion rate (affecting benzene and TCE).
- The RBCA and ROME defaults for exposure frequency and ingestion rate are higher than the GTS dataset, leading to an increase in dose. (N.B. ROME does not provide dose concentrations, the TCE value (threshold contaminant) was calculated by hand).
- Risc-Human, SFT, UMS and Vlier-humaan all have higher defaults for exposure frequency than the GTS dataset, and in addition Risc-Human and SFT have a higher default for ingestion rate. For all four systems, an increase in dose is observed, though it is only minor for UMS.

8.4.2. Inhalation of Indoor Air

Figure 26 presents the percentage change in Dose Concentrations for benzene and TCE for the Inhalation of Indoor Air from Soil exposure pathway. The findings are summarised below:

- The inhalation of indoor air dose for TCE predicted by RISC increases, which is explained by an increased exposure time and lung retention factor, however these effects are slightly off-set by a decrease in inhalation rate. The benzene dose decreases because the default exposure duration is lower which would heavily influence the dose output for any non-threshold contaminant.
- Risc-Human, Vlier-Humaan and UMS have a higher default exposure time than the GTS dataset, predicting increased inhalation of indoor air doses for both contaminants.
- The increase in dose observed in UMS and SFT is explained by the default for exposure frequency being higher than that used in the GTS.
- ROME does not provide dose concentrations, however for threshold substances (i.e. TCE) the dose can be calculated by hand. The observed decrease in TCE dose is attributable to the default inhalation rate being lower than the GTS value.
- The GTS and RBCA default exposure datasets are the same, consequently there is no variation.

8.4.3. Ingestion of Vegetables

Figure 26 presents the percentage change in Dose Concentrations for benzene and TCE for the Ingestion of Vegetables exposure pathway. The findings are summarised below:

- The decrease in dose concentration for RISC is attributable to the defaults for exposure duration and ingestion rate being lower.
- A decrease in dose is observed in Risc-Human, explained by a lower default ingestion rate than used in the GTS and a different fraction split between stem and root vegetables.
- SFT has a lower default ingestion rate, a higher default “fraction of vegetables consumed that are contaminated” and higher default exposure frequency, leading to an overall increase in dose concentration.
- UMS has a lower default exposure frequency but higher default “fraction of vegetables consumed that are contaminated”, giving an overall increase in dose.
- In Vlier-Humaan the fraction split between stem and root vegetables consumed varies. The uptake by the two vegetables types is not the same, leading to an overall increase in ingestion of vegetables dose when the default database is used.

8.5. Toxicological Input Parameters

The sensitivity testing of Toxicological inputs is focused on the variations produced in Quantified Risk Levels (QRLs) when the system default values are used instead of a GTS data set. The GTS and default toxicological input parameters are presented in Table 12 for benzene and TCE; thus accounting for a non-threshold and a threshold substance.

QRLs have been developed for both child and adult receptors based on the default toxicity data and the GTS data set for the following five exposure pathways:

1. Dermal Contact with Soil
2. Ingestion of Soil
3. Ingestion of Vegetables
4. Inhalation of Outdoor Dust
5. Inhalation of Indoor Air from impacted Soil.

Therefore for each pathway a risk levels has been calculated using the method employed in the country of origin and is compared to QRL derived from the GTS data set. With the exception of UMS, moving from the generic test site data set to the default toxicity data set requires the user to change only the TDD or the slope factor. In UMS an additional parameter is changed, the “resorption” factor, which is chemical and pathway specific.

The QRLs for the default toxicity data are compared to the QRLs for the GTS data, for TCE in Tables 13a to 13d, and for benzene in Tables 14a and 14b. The findings of the comparison are presented below.

- In RISC, RBCA and ROME the default toxicity data is very similar, if not the same, as the toxicity data defined for the GTS. Therefore, as expected there is minimal change in the QRLs.
- In Risc-Human and Vlier-Humaan the GTS toxicity data and the default toxicity data vary significantly, hence the difference in the QRLs between the two data sets.
- In UMS, the TCE default toxicity data is higher than the GTS toxicity data, resulting in a lower QRL.

- In UMS, the benzene default TDD is very similar to the GTS TDD, however, for the dermal contact pathway, the default resorption factor (3.5%) is less than the GTS resorption factor (100%), therefore the QRL from this pathway is reduced for the default data set.

The results of the Toxicological Sensitivity testing are relatively straight forward, however, the procedure highlights the difficulties of comparing the methods of QRL calculation and the different ways that the output is presented to the risk assessor.

9. PHASE IV – CASE STUDY TEST SITES

9.1. Introduction

Phase IV of this Risk Comparisons Study comprises the use of real site data to produce outputs for comparison.

Information for five Case Study Test Sites were provided by the sponsors of the project. Each site was reviewed and the likely critical pathways for assessment were defined as the conceptual site model for each site. The risk assessments carried out on each of the test sites was not intended to be a comprehensive risk assessment for any legal purposes, moreover, the intention was to explore further the findings made during the Phase III testing by isolating specific source pathway receptor linkages.

Test Site 1 was a former industrial site located in a mixed commercial, industrial and residential setting. The site was included as the main issue at the site surrounds impacts to groundwater and off-site groundwater migration. The site and surrounding area has been extensively investigated and the plume of impact is well defined both on-site and down hydraulic gradient, allowing the opportunity to use the risk assessment systems in a predictive capacity.

Test Site 2 is part of a former gas works where all of the on-site structures associated with production have been decommissioned and demolished. Residual impacts have been identified in the soils typical of the former land-use. Therefore the pathways of interest at this site comprise direct contact exposure pathways.

Test Site 3 is a Pulverised Fuel Ash Landfill which has been extensively investigated. However its current use is as agricultural land, therefore this site has been assessed from the perspective of potential re-development for residential end-use. The impacts identified at the site are again present within the soils and therefore the pathways of interest at this site comprise direct contact exposure pathways.

Test Site 4 is an active agro-chemicals manufacturing plant and therefore the exposure scenarios include on-site worker exposures. The impacts identified comprise volatile organic compounds and as such the focus for this site is on the assessment of the indoor air pathway.

Test Site 5 is a former petrol/gasoline filling station with typical soil and groundwater impacts. The focus of the assessment at this site is on the indoor air pathway with the site offering the added benefit of measured indoor air concentrations of benzene allowing, albeit in limited form, the opportunity to compare predicted concentrations with measured values.

To summarise the scope of work for Phase IV, Table 15 presents a list of the active pathways considered at each of the five test sites and Table 16 presents a matrix of active pathway against risk assessment system for the five test sites.

Each of the following five sections is devoted to a single test site and comprises:

- A summary of the site and testing schedule;
- Background information about the site;
- Conceptual site model definition,
- Results and Discussion.

10. TEST SITE 1 - FORMER LUBE OIL PLANT

10.1. Site Summary

This test site comprises a source of trichloroethene (TCE) released to soil and groundwater at a former lube oil manufacturing site, with a well-defined plume of TCE and TCE degradation products (as *cis*-1,2 dichloroethene (cDCE) and vinyl chloride (VC)) in groundwater at distances of up to 600m downgradient of the site. This test site provided a good dataset for the comparison of those risk systems that include algorithms/models for the prediction of contaminant fate and transport in groundwater.

The following risk assessment systems were run for this test site:

JAGG • P20-RTW • RBCA • ROME • RISC

10.2. Background

The site, an industrial works closed in 1998, is located in an area of commercial, industrial and residential use, and was active for a number of years. The area of the site is approximately 10,000 m². The soil and groundwater beneath the site is impacted with chlorinated solvents. A groundwater plume comprising TCE, cDCE and VC, extends from the site boundary to a point 625m to the north, where a large canal intercepts the groundwater.

The groundwater plume extends beneath a residential area immediately down hydraulic gradient of the site.

Two forms of remedial action have been undertaken at the site to address the chlorinated hydrocarbon plume. Firstly, a Soil Vapour Extraction (SVE) system was installed in April 1999 and secondly, a 'pump and treat' groundwater remediation system was installed in June 1999. For this risk assessment, Arcadis has used the available groundwater monitoring data prior to the remedial actions.

A simplified site schematic is presented as Figure 27 illustrating the locations of investigation borings, groundwater monitoring wells and a simplified geological cross section.

10.3. Conceptual Model

The assessment of Test Site 1 comprises the prediction of receptor point concentrations for groundwater at three distances from the site (57.5m, 120m and 625m) derived from a groundwater source. Indoor air pathways are also considered as part of this assessment, for a residential property located 120m down-gradient of the site. The following sections describe the definition of the sources, pathways and receptors considered for Test Site 1.

10.3.1. Sources

The selected Contaminants of Concern (COCs) are:

- Trichloroethene (TCE)
- *cis*-1,2-Dichloroethene (cDCE)
- Vinyl Chloride (VC)

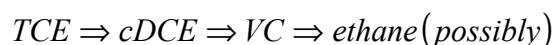
Groundwater monitoring data indicate that the source is located in the vicinity of monitoring wells GMW4 and GMW5 (see Figure 27). No soil concentration data were provided, therefore only a groundwater source has been modelled.

The concentration of TCE is elevated in groundwater beneath the site, but decreases rapidly with distance from the site boundary. At a distance of approximately 200m north of the site boundary, the concentration of TCE is negligible.

Dissolved oxygen data were provided for groundwater in monitor wells at the site and in the areas in and outside the plume down gradient of the site. In general, dissolved oxygen concentrations are well below oxygen solubility, with similar oxygen concentrations inside the TCE plume and outside the plume. It is tentatively concluded that reductive de-chlorination is facilitated by naturally low redox potentials and low dissolved oxygen concentrations in the aquifer.

It is considered that TCE is degrading by reductive de-chlorination under anaerobic conditions, to form cDCE. The cDCE concentration in groundwater increases correspondingly with the decrease in TCE. However, from a distance of 200m north of the site boundary to the canal edge, 625m to the north, the cDCE concentration is the main COC indicating that the de-chlorination reaction has stalled or slowed. Low concentrations of VC have been recorded close to the canal, indicating that cDCE is probably also degrading by reductive de-chlorination. There are no measurements of ethene and therefore no direct evidence of the degradation of VC.

In general terms it is assumed that there is conservation of *equivalent* contaminant mass within the plume, as the contaminant transforms:



In this context, equivalent mass refers to the mass of carbon because this is conserved through the above transformation process. Net mass will be lost due to the step-wise loss of chlorine and substitution with hydrogen. However, two mechanisms will act to reduce contaminant mass:

- volatilisation to the unsaturated zone, especially for VC,
- adsorption to the aquifer matrix.

10.3.2. Pathways

The local geology comprises sands with localised silts and clays overlying gravelly sands and gravels. This in turn overlies clay, interbedded with sand, at a depth of between 12 to 15m below ground level (bgl).

The geology beneath the site is permeable, comprising predominantly sands and gravels. Site aquifer test data indicate a hydraulic conductivity of 40m/d. Effective porosity is estimated at 0.2 (20%). The hydraulic gradient measured from groundwater elevation contours is 0.009.

Therefore the groundwater flow velocity is estimated to be 1.8m/day from:

$$v = \frac{K i}{n}$$

where:

v= groundwater velocity (m/d)

k= hydraulic conductivity (m/d)

i = hydraulic gradient (-)

n= effective porosity (-)

These data were used to model the migration of groundwater through a sandy gravel media.

10.3.3. Receptors

A number of residential properties lie to the north of the site, and are considered active receptors to the volatilisation of chlorinated hydrocarbons from the plume located beneath the properties. Neighbouring residents have been included in the modelling process, with inhalation of vapours in indoor air (originating from a groundwater source) being the active exposure pathway.

10.3.4. Input Data

Table 17 lists the key physical input parameters used for this test site.

10.3.5. Output

Receptor point concentrations are provided by all models used for this test site.

Table 22 lists the measured groundwater concentrations in the off site plume that were used as a benchmark against which to compare model output. Receptor point concentrations were output from the models at distances corresponding to monitor wells with measured contaminant concentrations, these being GWM8 and GWM9 at 57.5m and 120m away from the source respectively.

JAGG is not structured to allow the user to output a concentration at a user-defined distance. The concentration of contaminants is calculated for a point at a distance from the pollution source, which corresponds to one year of groundwater flow, through a maximum of 100m. The system specifies that the groundwater quality criteria must be met at this point. Furthermore, the concentration is only reported at 10m intervals up to the maximum distance of 100m. In this study the concentration reported by JAGG at 60m is assumed to represent the predicted concentration at 57.5m, resulting in a small over-prediction.

Only one of the systems used, RISC, has the capability to model the transport of COCs to indoor air, and from this derive estimates of dose and risk.

10.3.6. Scenarios

The predicted receptor point concentrations were compared with measured concentrations in monitor wells GWM8 and GWM9. The measured concentrations were used as a reference point only, there being no attempt to calibrate the models to these measured values.

Two scenarios were run, as outlined below:

- **Scenario 1: TCE as the only COC.** Degradation was included, and the receptor point was located within 200m of the site boundary i.e. within the observed limits of the TCE plume.
- **Scenario 2: The sum of TCE, cDCE and VC as a single COC.** Degradation was not included and the receptor point was located within 625m of the site boundary. cDCE or VC were not considered as independent COCs as their concentrations could potentially increase as a result of the TCE degradation process.

In the second scenario the chemical properties of TCE, DCE and VC are assumed to be similar, however in reality, each of these substances have distinctly different fate and transport properties. Relative to TCE, cDCE is more soluble and has a lower organic carbon partition coefficient, and is generally less readily degraded. Therefore cDCE is generally more mobile than TCE. Relative to TCE, VC is more volatile and will therefore partition more readily to the soil gas phase in the unsaturated zone.

10.4. Results and Discussion

The output results are presented in Table 18 and a discussion of the results is presented in the following sections for both scenarios outlined above.

10.4.1. Receptor Point Concentrations

Scenario 1

The receptor point concentrations of TCE in groundwater at 57.5m and 120m distances from the source are presented in Figure 28(a). The figure shows that all of the systems evaluated predict concentrations within the same order of magnitude as the measured concentration at a distance of 57.5m. All the systems produce over estimates, with the over estimation effect appearing to increase with distance. P20, RBCA Toolkit and RISC produce values in a narrow range, which is to be expected given that they use very similar transport equations. The JAGG system predicts a concentration very close to the measured concentration, even though the receptor point is automatically set at 60m (as opposed to the 57.5m used in the scenario).

With reference to the GTS results, it was found that the groundwater concentrations predicted by JAGG were also lower than those predicted by the other systems. It was determined that JAGG produced lower predicted groundwater concentrations than P20, RBCA Toolkit and RISC because of the different method of calculation used in JAGG (Section 7.3.2).

The dispersivity is hard-wired in the JAGG code and assumed to be distance-dependant. This is reported to be based on the work by Gelhar *et al.* but is not referenced in the manual. However, it is clear that JAGG uses dispersivity values that are lower than that specified for this test site. A lower dispersivity would, in general, be expected to produce *higher* concentrations.

The transport time used by JAGG is limited to 1 year, whereas RBCA and P20 are steady state models and RISC was run to 50 years. It was notable that in the RISC transient simulation maximum concentrations were reached after 2 years. It is likely that the JAGG receptor point concentrations are lower than the other system outputs because the concentrations are increasing towards the final steady state values i.e. a transport time of only one year does not allow concentrations to reach maximum values.

It is likely that the significant over-prediction that occurred in estimating the TCE concentrations is due to the difficulty in estimating the TCE degradation rate rather than an inadequacy in the systems.

Scenario 2

In this scenario the system outputs comprised a predicted TCE concentration and were compared to the sum of the measured concentrations of TCE, cDCE and VC. The combining of the solvent substances was intended to overcome the limitations imposed on the comparison of modelled to measured concentrations where degradation of TCE is involved.

Receptor point concentrations of “TCE” are presented on Figure 28(b) and discussion of the results is presented below.

All the systems evaluated predict concentrations within the same order of magnitude as the measured concentration, with the systems predicting a concentration approximately 30% higher than the measured concentration at 120m. Thereby being both reasonably accurate whilst also being somewhat conservative.

No measured concentration was available at 625m, although a canal is located at this distance and therefore, the measured concentration was replaced with a water quality compliance standard, which has been taken to be 0.01mg/l TCE. All of the systems predict that the concentration at the surface water feature would be greater than this water quality standard.

10.4.2. Dose Prediction

Of the risk systems run for this test site doses can only be calculated by RISC. Therefore, no comparison of doses produced by the systems was feasible for this test site.

11. TEST SITE 2 - FORMER GASWORKS

11.1. Summary

This test site comprises part of a former gas works site with benzo(a)pyrene, cadmium, free cyanides and naphthalene measured within the shallow soil. The test site allows comparison of direct contact pathways commonly assessed for residential land-use; ingestion of soil, dermal contact with soil and ingestion of vegetables.

The following risk assessment systems were run for this test site:

RBCA • RISC • Risc-Human • SFT99:06 • UMS • Vlier-Humaan • CLEA

11.2. Background

This test site formed part of a gas works site. The gas holders and purifiers were demolished before 1990, however the work sheds, gas works house and former governor house remain on site. The area defined as Test Site 2 is a two storey residential property (gas works house) and associated single storey out-building to the south of the site with garden. It is assumed that the end-use of the site will remain residential.

A site layout plan is presented as Figure 29.

The test site has been investigated with boreholes, and impacts have been identified with contaminants characteristic of a former gas works, including:

- ammonium
- cyanide
- PAHs presumably derived from coal tar
- metals.

11.3. Conceptual Model

Outlined below are the active source-pathway-receptor linkages to be assessed as part of the modelling process.

Sufficient information was not available to assess the risks to the minor aquifer beneath the site, as might normally be undertaken in a full environmental and human health risk assessment.

Volatilisation to indoor air from VOCs and ammonium may also be an active pathway for this site. However the identified organic contaminants are relatively non-volatile, whilst the fate and transport of ammonium is dependent on pH cation exchange, which the majority of systems cannot model. Therefore volatilization pathways have not been considered.

11.3.1. Sources

The assessment has been carried out for the following Contaminants of Concern (COCs):

1. Benzo(a)pyrene (non-threshold)
2. Cadmium (threshold)
3. Free Cyanide (threshold)
4. Naphthalene (threshold)

11.3.2. Pathways

The following exposure pathways are considered to be active:

- Direct ingestion of soil by residential receptor (adult)
- Dermal contact with soil by residential receptor (adult)
- Vegetable ingestion by residential receptor (adult)

11.3.3. Receptors

The key receptors for this site are the residential occupants of the building. For the risk assessment it has been assumed that there will be a family present on the site.

11.3.4. Input Data

The parameter values used to define the conceptual model are presented in Table 19.

11.3.5. Summary of Assumed Active Exposure Pathways

For this site, the exposure of residential receptors to benzo(a)pyrene, cadmium, free cyanide and naphthalene has been evaluated, via direct ingestion, dermal contact and vegetable ingestion.

11.4. Results and Discussion

Table 20 presents the dose concentrations produced by the systems included in the evaluation of this site. Figures 30A to 30D graphically represent the dose concentrations for benzo(a)pyrene, cadmium, free cyanide and naphthalene respectively.

Predicted dose concentrations are presented for each substance based on the default chemical, toxicological and exposure parameters in the database of each system. This has particular consequences for non-threshold substances such as benzo(a)pyrene where systems may incorporate different averaging approaches, effectively rendering the results non-comparable. Cadmium, free cyanides and naphthalene are all treated as threshold substances, for which the exposure duration is assumed to equal the averaging time in all systems.

Ingestion of Soil

For this pathway the chemical properties of a substance are not incorporated into the algorithms for calculating the dose, with the exception of whether the chemical is a non-threshold or a threshold substance in which case the averaging time may vary.

Cadmium, Free Cyanide and Naphthalene

The results of the sensitivity analysis carried out on the exposure data (Figure 26) indicates that a wide variation in results can be expected when the default exposure data are used for each system.

- SFT and Risc-Human predict identical doses, as would be expected from the results of the GTS and the sensitivity analysis on the exposure data.
- The dose predicted by RBCA is approximately double that predicted by SFT and Risc-Human, the main reason for which is that the default soil ingestion rate is approximately double that of SFT and Risc-Human.
- The doses predicted by UMS are less than that of SFT and Risc-Human, again explained by variations in exposure data, with the exposure frequency and ingestion rate both being less than that of SFT and Risc-Human.
- UMS also incorporates a chemical-specific “resorption” factor to account for the difference between intake and uptake by the body. This is set at 5%, 70% and 85% for cadmium, free cyanides and naphthalene respectively, hence explaining why the cadmium doses differs from SFT and Risc-Human significantly, while the free cyanide and naphthalene doses differ by smaller factors.
- Vlier-Humaan predicts a lower dose than Risc-Human due to a lower default ingestion rate and a more complex algorithm for modelling the exposure frequency.
- CLEA predicts a greater dose than any of the other systems. CLEA is a probabilistic model and in keeping with the way in which the system is intended to be used in its country of origin the results quoted for this test site are for the 95th percentile. However for Free Cyanide the predicted dose concentrations are similar to RBCA.
- RISC produces a lower dose than the other systems due to the variation in exposure data. The RISC system is unique in providing two sets of default exposure parameters, based on a “typical” exposure and a “reasonable maximum exposure”. For this test site, the “typical” values were selected. Selecting the “reasonable maximum exposure” data would result in doses more similar to those of RBCA, as seen in the GTS.

Benzo(a)pyrene

The main differences seen for this chemical relate to differences in the approaches adopted for defining an averaging time for non-threshold compounds:

- SFT, Risc-Human and Vlier-Humaan all assume that the averaging time is equal to the exposure duration for an “adult” dose. The Vlier-Humaan dose is lower because of the hard wiring as discussed in Section 7.1.1.
- The UMS dose is more similar to SFT and Risc-Human than might be expected due to the differences in exposure data. However, for non-threshold compounds UMS bases the non-threshold dose on a lifetime receptor, in effect summing the dose for an adult and the dose for a child.
- RBCA predicts a lower dose than SFT and Risc-Human in comparison to the differences seen for the threshold compounds. This is because RBCA predicts a dose averaged over a lifetime for non-threshold compounds.
- RISC and CLEA predict respectively lower and higher doses based on the selected exposure data though they vary in their treatment of non-threshold compounds. RISC predicts a lifetime averaged dose where CLEA defines the averaging time to be equal to the exposure duration (dose is for age class 17).

Dermal Contact with Soil

In addition to the variation in default exposure parameters, the variation in default chemical properties also helps to explain some of the differences seen between the systems for this pathway.

Cadmium

RISC and SFT are the only two systems that predict a dose for cadmium. The main explanation behind the differences between the predicted doses for these two systems is the value assigned to the chemical specific absorption factor in the default chemical databases. The SFT value is 140 times greater than the RISC value. There is also a small difference in the default exposure data sets defined for this pathway.

The remaining systems do not predict doses for cadmium as it is assumed that there is no mechanism by which metals are taken into the body via dermal contact.

Naphthalene and Free Cyanide

- Cyanide is not assigned soil-skin dermal absorption properties in RBCA or Risc-Human therefore no dose is presented.
- SFT predicts doses for naphthalene and free cyanide comparable to Vlier-Humaan, while in the GTS presents results greater than Vlier-Humaan. This can be attributed to a combination of the default exposure parameters and default chemical properties in both systems providing similar doses, even though the algorithms in the systems differ in form.
- Variation in the doses predicted by UMS in comparison to SFT and Vlier-Humaan mainly occur due to differences in the chemical specific “resorption” factors.

Benzo (a)pyrene

The dose for benzo(a)pyrene from CLEA is the greatest; in keeping with the national guidance the 95th percentile dose is presented.

RBCA predicts doses greater than RISC while in the GTS the predicted doses are identical. In RISC the “typical” exposure scenario has been selected – the “reasonable maximum exposure” scenario would provide doses similar to RBCA.

- UMS the doses for benzo(a)pyrene are greater than SFT while they are similar in the GTS. A lifetime dose is presented for UMS, which includes exposure as a child and an adult.
- SFT predicts doses for benzo(a)pyrene comparable to Vlier-Humaan, while in the GTS presents results greater than Vlier-Humaan. This can be attributed to a combination of the default exposure parameters and default chemical properties in both systems providing similar doses.

Ingestion of Vegetables

The differences seen for this pathway are dependent on exposure data but also on the definition of Bio-Concentration Factors (BCFs), for which the default values vary considerably between the systems.

Benzo(a)pyrene

The explanations behind the differences seen for this pathway are partly due to the differences in the definition of the averaging time, as explained for the ingestion of soil. However, there are some other key differences, as explained below:

- SFT predicts a much greater dose than any of the other systems, due to the default BCF.
- Vlier-Humaan predicts a slightly higher dose than Risc-Human, contrary to the results of the GTS. This difference is explained by the greater exposure predicted by the Vlier-Humaan default data, as seen in Figure 26.

- RISC produces a lower dose than the other systems. This trend was seen in the GTS results, with the exception of UMS, which produced a lower dose. The RISC system is unique in providing two sets of default exposure parameters, based on a “typical” exposure and a “reasonable maximum exposure”. For this test site, the “typical” values were selected. Selecting the “reasonable maximum exposure” data would result in a higher dose.
- CLEA predicts a higher dose than the majority of systems in comparison to the GTS. This difference is due to the results being presented for the 95th percentile resulting in the dose being based on different exposure data.

Cadmium

- The dose predicted by Risc-Human is greater than that predicted by UMS, as would be expected from both the results of the GTS and the sensitivity testing of the chemical input parameters presented on Figure 24.
- The dose predicted by Vlier-Humaan is greater than that predicted by Risc-Human where both the sensitivity analysis for the chemical input parameters and the results of the GTS suggest that we would expect to see the Vlier-Humaan dose to be less than the Risc-Human dose. However, referring back to the sensitivity analysis for the exposure data it can be seen that the default exposure data set for Vlier-Humaan results in a higher predicted dose than for Risc-Human. Therefore it is the default exposure data that are controlling the relationship between Vlier-Humaan and Risc-Human for this pathway.
- The dose predicted by RISC is now very similar to that predicted by Risc-Human where it was greater in the results of the GTS. The sensitivity analysis of the chemical input parameters (Figure 24) indicates that using the default chemical data for cadmium would result in a comparatively higher dose from Risc-Human than from RISC, which in this case leads to the Risc-Human dose now being similar to the RISC dose.
- The dose predicted by CLEA is greater than that of other systems, as was seen in the GTS. Default chemical data may now be contributing to that difference though the probabilistic modelling of this pathway in CLEA makes it difficult to refine the explanation behind the differences seen here. The 95th percentile is presented in the test site where the 50th percentile was presented in the GTS.

Free Cyanide

- Four of the systems provide a dose for vegetables contaminated with free cyanide. There is a large variation between the predicted doses, predominantly as a result of different BCFs.

Naphthalene

Many of the differences seen here are as a result of the definition of the BCF.

- SFT and Risc-Human use similar BCFs (calculated in Risc-Human) that are much greater than the BCFs defined in Vlier-Humaan. The remaining difference between SFT and Risc-Human is due to differences in the default exposure data.
- UMS predicts a lower dose than SFT and Risc-Human. This may be due to the value of the calculated BCF, however, this value is not given as an output of the UMS software and therefore cannot be verified.
- RISC predicts a lower dose than SFT and Risc-Human, as a result of the difference in calculation of the BCF.
- The dose predicted by CLEA lies between that of Risc-Human and Vlier-Humaan. Default chemical data may now be contributing to that difference though the probabilistic modelling of this pathway in CLEA makes it difficult to refine the explanation behind the differences seen here.

12. TEST SITE 3 - FLY ASH LANDFILL

12.1. Summary

This test site is a former landfill site used for disposal of pulverised fuel ash (PFA) from a nearby coal fired power station. The landfill operated as a number of separate disposal cells. PFA was transported to site in a slurry form and was then dewatered *in situ* in settling lagoons. A site layout plan is presented on Figure 31.

Prior to use as a landfill, the site was a sand and gravel quarry, where the aggregates were extracted from the complete thickness of a sand and gravel unit, known as River Terrace deposits. It is understood that aggregate extraction has removed the majority of the sand and gravel that previously formed the minor aquifer and therefore it is considered that the aquifer has been largely removed within and outside the site. The site has since been filled with PFA and restored to original ground level using stockpiled soil and the deposited PFA is now partly below and partly above the water table. The surface is now vegetated with grass and used for grazing livestock.

The site is located in a low-lying river valley adjacent to a large river. Groundwater is present at a shallow depth, generally about 1m and consequently there are numerous surface water features on-site and surrounding the site. These surface water features are present as the impounded aggregate quarries that have been restored as amenity lakes that may be used for fishing, wildlife habitats and/or water sports.

This test site provided a number of opportunities for evaluating model performance, including certain active exposure pathways, as follows:

- General ability of the risk models to evaluate landfill sites
- Evaluation of soil exposure and vegetable uptake models for inorganic substances
- Evaluation of the dust exposure models for inorganic substances (PFA is a fine grained material readily mobilised into the atmosphere)

A number of other exposure pathways are also potentially active at this site although these were not considered for various reasons:

Groundwater monitoring data was available from monitor wells on the site boundary. However, groundwater exposure assessments were not evaluated due to the difficulty in establishing groundwater flow and transport pathways in the complex residual aquifer system.

An indirect human exposure pathway is active at this site via grazing livestock, which may consume PFA and PFA-derived contaminants through soil and vegetable ingestion. However, relatively few models can evaluate this pathway and therefore there was no opportunity of comparative assessment.

The following risk models were run for this test site:

RBCA • RISC • Risc-Human • SFT99:06 • UMS • Vlier-Humaan • CLEA

12.2. Conceptual Model

The assessment of Test Site 3 comprises the prediction of dose concentrations from impacted superficial soils. The following sections describe the definition of the sources, pathways and receptors that comprise the conceptual site model for the assessment of Test Site 3.

At present the site is used for agricultural purposes, however for this assessment it has been assumed that the risk assessment is being carried out for a proposed residential re-development.

12.2.1. Sources

Based on a review of the available site data, the following contaminants of concern were considered to be present and of relevance to the pathways defined for the assessment:

- Arsenic
- Cadmium
- Chromium VI

12.2.2. Pathways

The local geology comprises alluvial silts, clays, sands and sands and gravels, underlain by marl and mudstones.

A minor aquifer, the River Terrace sands and gravels, was formerly present beneath and adjacent to the site. The aquifer has been extracted at the site to form the pits in which the PFA was deposited. Other worked-out pits are present adjacent to the site. Some in situ River Terrace may be present locally, largely beneath access roads and buildings, forming a residual aquifer of complex shape.

The site is underlain at depth by the Sherwood Sandstone, a major aquifer, separated from the River Terrace and in filled pits by the Mercia Mudstone which is described as a non-aquifer. The major aquifer is considered to be too deep and concealed beneath the mudstone to be significantly affected by site-derived contaminants.

The following exposure pathways are considered to be active:

- Direct ingestion of soil by residential receptor
- Dermal contact with soil by residential receptor
- Vegetable ingestion by residential receptor
- Inhalation of dust

12.2.3. Receptors

The site is currently used as grazing land, although for the purpose of the risk assessment trial, it has been assumed that a residential property will be located at the site as a hypothetical re-development project. Therefore, the assumed site occupants include children and adults.

It should be noted that only Risc-Human and Vlier-Humaan have the capability to assess exposure via ingestion of meat from animals that have grazed on contaminated sites.

12.2.4. Input Data

The parameter values used to define the conceptual model are presented in Table 21.

Model default values were used for chemical properties and for exposure parameters. Site-specific data were used for all physical and spatial properties.

12.2.5. Summary of Assumed Active Exposure Pathways

For this site, the exposure of residential receptors to arsenic, cadmium and chromium VI has been evaluated, via direct ingestion, dermal contact, vegetable ingestion and inhalation of dust.

12.3. Results and Discussion

Predicted dose concentrations and quantified risk levels are presented in Table 22.

Dose concentrations are presented for ingestion of soil, dermal contact with soil, ingestion of vegetables and inhalation of dust with associated risk levels as presented in Figures 32A to 32C, for arsenic, cadmium and chromium VI respectively. Four graphs are presented on each figure, one for each pathway.

Dose concentrations have been presented for arsenic, cadmium and chromium VI as both threshold and non-threshold substances, depending on how these elements are treated in each system:

- Arsenic was treated as a threshold and as a non-threshold substance by most models, except CLEA, which treated it as a non-threshold substance only.
- Cadmium was treated as a threshold substance by all models.
- Chromium VI was generally treated as a threshold contaminant by most models, except for the inhalation pathway where it was treated as both. Vlier-humaan treated the substance as a non-threshold substance for all pathways.

ROME does not present dose concentrations as an output therefore only the risks associated with these pathways are reported.

12.3.1. Ingestion of soil

For this pathway the chemical properties of the substances are not incorporated into the algorithms for calculating the dose, with the exception of whether the chemical is a non-threshold or a threshold substance in which case the averaging time may vary.

- SFT and Risc-Human predict similar doses for all contaminants, as would be expected from the results of the GTS and the sensitivity analysis on the exposure data.
- The dose predicted by RBCA is approximately double that predicted by SFT and Risc-Human, the main reason for which is that the default soil ingestion rate is approximately double that of SFT and Risc-Human.
- The doses predicted by UMS are less than that of SFT and Risc-Human, again explained by variations in exposure data, with the exposure frequency and ingestion rate both being less than that of SFT and Risc-Human.

- UMS also incorporates a chemical-specific “resorption” factor to account for the difference between intake and uptake by the body. This is set at 5%, 100% and 5% for cadmium, arsenic and chromium VI respectively, hence explaining why the cadmium and chromium VI doses differ from SFT and Risc-Human significantly, while the arsenic dose differs by a smaller factor.
- Vlier-Humaan predicts a lower dose than Risc-Human due to a lower default ingestion rate and a more complex algorithm for modelling the exposure frequency.
- CLEA predicts a greater dose than any of the other systems. CLEA is a probabilistic model and in keeping with the way in which the system is intended to be used in its country of origin the results quoted for this test site are for the 95th percentile.
- RISC produces lower doses than the majority of other systems due to the variation in exposure data. The RISC system is unique in providing two sets of default exposure parameters, based on a “typical” exposure and a “reasonable maximum exposure”. For this test site, the “typical” values were selected. Selecting the “reasonable maximum exposure” data would result in doses more similar to those of RBCA, as seen in the GTS.

12.3.2. Dermal Contact with Soil

In addition to the expected variation as a result of using the default exposure parameters, the variation in default chemical properties are a major contributing factor to the variations identified here.

Cadmium

- RISC and SFT are the only two systems that predict a dose for cadmium. The main explanation behind the differences between the predicted doses for these two systems is the value assigned to the chemical specific absorption factor in the default chemical databases. The SFT value is 140 times greater than the RISC value. There is also a small difference in the default exposure data sets defined for this pathway.
- The remaining systems do not predict doses for cadmium as it is assumed that there is no mechanism by which metals are taken into the body via dermal contact.

Arsenic and Chromium

- UMS predicts the greatest dose for chromium VI. The chemical specific dermal absorption factor is set at 0.7 in UMS, compared with less than 0.1 in SFT and RISC. The absorption factor is set at zero for arsenic therefore no dose is produced for UMS.
- The dermal absorption factor is the same for arsenic in RISC and SFT, therefore the differences seen between the predicted doses are explained by the use of the “typical” exposure dataset in RISC.

12.3.3. Ingestion of vegetables

The differences seen for this pathway are largely dependent on exposure data but also on the definition of Bio-Concentration Factors (BCFs), for which the default values vary considerably between the systems.

Cadmium and Arsenic

- The dose predicted by Risc-Human is greater than that predicted by UMS for cadmium, as would be expected from both the results of the GTS and the sensitivity testing of the chemical input parameters (See Figure 24). The UMS dose for arsenic is large in comparison to Risc-Human due to a variation in default chemical specific properties.

- The dose predicted by Vlier-Humaan is greater than that predicted by Risc-Human where both the sensitivity analysis for the chemical input parameters and the results of the GTS suggest that we may expect to see the Vlier-Humaan dose to be less than the Risc-Human dose. However, referring back to the sensitivity analysis for the exposure data it can be seen that the default exposure data set for Vlier-Humaan results in a higher predicted dose than for Risc-Human. Therefore it is the default exposure data that are defining the relationship between the results produced by Vlier-Humaan and Risc-Human for this pathway.
- The dose predicted by RISC is very similar to that predicted by Risc-Human where it was greater in the results of the GTS. The sensitivity analysis of the chemical input parameters (Figure 24) indicates that using the default chemical data for cadmium would result in a comparatively higher dose from Risc-Human than from RISC, which in this case leads to the Risc-Human dose being similar to the RISC dose. This is the case for arsenic as well as cadmium.
- The dose predicted by CLEA lies between that of Risc-Human and Vlier-Humaan. Default chemical data may be contributing to that difference though the probabilistic modelling of this pathway in CLEA makes it difficult to refine the explanation behind the differences seen here.

Chromium VI

- The relationship between Risc-Human, UMS and Vlier-Human is the same as for cadmium, however the dose predicted by CLEA is greater. This difference is due to the results being presented for the 95th percentile resulting in the dose being based on different exposure data, combined with the chemical specific properties for chromium incorporated into the system.

12.3.4. Inhalation of dust outdoor

The difference between the predicted doses is similar for all three compounds. The dose produced by Vlier-Humaan is significantly greater than the remaining systems, due to a slightly higher default dust concentration in air, but more significantly due to a larger exposure duration in comparison to the remaining systems.

The remaining differences between the systems is as expected, given the variation in dust inhalation models outlined in the GTS.

13. TEST SITE 4 - ACTIVE CHEMICAL MANUFACTURING SITE

13.1. Introduction

This test site comprises an active agro-chemicals production site. The site began activity in 1938 and has manufactured numerous chemicals ranging from chlorines and chlorinated solvents to pesticides and quinones. The site is located on the banks of a major river; the hydrological relationship between the site and the river has been assessed and there is a seasonal variation of losing and gaining.

Area-wide site investigations exist for 1989 and 1990 together with a more recent borehole investigation of the north-east corner of the site undertaken in 1997. Extensive groundwater monitoring has been completed. The north-eastern area of the site provides the focus of the risk assessment and a site layout plan of this part of the site is presented on Figure 33.

Shallow and deep soil sources have been identified, comprising by tetrachloroethene (PCE), trichloroethene (TCE) and lindane (*g*-HCH - *gamma*-Hexachlorocyclohexane). A dissolved plume of extends over a lateral area of approximately 14,000m². This plume originates from areas VII and VIII of the site as shown on Figure 33. Groundwater rests at approximately 9m below ground level.

The test site allows comparison of direct contact pathways for commercial workers; firstly with the shallow soil source and secondly via the inhalation of indoor air impacted by vapours originating from the deep soil source.

The following risk assessment systems were run for this test site:

RBCA • RISC • Risc-Human • JAGG • SFT99:06 • UMS • Vlier-Humaan

13.2. Conceptual Model

The assessment of Test Site 4 comprises the prediction of indoor air concentrations from the identified impacted soils and the development of dose concentrations to assess that pathway, and also the calculation of dose concentrations for the potential direct exposure pathways with the impacted superficial soils. The following sections describe the definition of the sources, pathways and receptors that comprise the conceptual site model for the assessment of Test Site 4.

13.2.1. Sources

The Contaminants of Concern (COCs) are assumed to be:

- Tetrachloroethene (PCE)
- Trichloroethene (TCE)
- Lindane

13.2.2. Pathways

The following exposure pathways are considered to be active:

- Volatilisation from a deep soil source on site and exposure of commercial workers via inhalation of indoor air.
- Direct exposure of on-site commercial workers from a soil source, including incidental dermal contact and incidental ingestion.

13.2.3. Receptors

The river has not been considered as a receptor due to its seasonal fluctuations, which cannot be modelled effectively in any of the systems. The only active receptors are considered to be the commercial workers on site.

It should be noted that Risc-Human and SFT99:06 have the capability to model a commercial worker receptor, however neither system has default exposure parameters to account for a commercial land-use. Therefore the default adult parameters were used where no other exposure parameters were provided.

13.2.4. Input Data

The parameter values used to define the conceptual model are presented in Table 23.

Where site-specific data were not available, system default values have been adopted to simulate how each system may be used to assess a site. In the case of this site default values have been used for chemical properties and for exposure parameters.

Lindane is not included in the JAGG chemical database and no new chemicals can be added, therefore results were not produced by this risk assessment systems for Lindane.

13.2.5. Summary of Assumed Active Exposure Pathways

For this site the exposure of commercial workers to PCE, TCE and Lindane has been evaluated, via ingestion and dermal contact with shallow soil and inhalation of indoor air impacted by vapours originating from a deep soil source.

13.3. Results and Discussion

The predicted receptor point concentrations in indoor air from a deep soil source are reported for all three contaminants of concern. Dose concentrations are presented for ingestion of soil, dermal contact with soil and inhalation of indoor air.

13.3.1. Receptor Point Concentrations

The concentrations of indoor air are reported in Table 24 and presented on Figure 34A.

- As would be expected from the results of the GTS the concentrations of all contaminants of concern in indoor air predicted by UMS are greater than that predicted by any of the other systems.

- RBCA, RISC and Risc-Human predict results that are more similar for Lindane, than for TCE and PCE. This finding is explained by similarities and differences in the chemical properties of Lindane, TCE and PCE between models. Model defaults were generally retained for chemical properties, however, as a less common contaminant Lindane was not present in all of the databases. Where this was the case the properties of Lindane were taken from Risc-Human.
- Vlier-Humaan predicts lower concentrations than Risc-Human, partly explained by the results of the GTS and partly due to the difference in the definition of building parameters in Vlier-Humaan: the footprint of the building is hard-wired to be smaller than that in Risc-Human but the height is an input parameter, resulting smaller indoor air concentrations.

13.3.2. Doses

The predicted dose concentrations are presented in Figures 34B to 34D for PCE, TCE and Lindane respectively. Three graphs are presented on each figure, one for each pathway.

Ingestion of soil

All of the systems predict doses based on the contaminants being treated as threshold compounds, i.e. a risk level derived from this dose describes the compound as being a threshold compound and the averaging time is equal to the exposure duration in the dose calculation. Differences are therefore described on this basis, though it should be noted that a number of systems also consider them as non-threshold compounds where the dose may be averaged over a lifetime.

- The results for TCE and PCE are identical across all models for the ingestion of soil. This is expected since the only chemical property required for predicting these doses (source concentration), is defined to be equal for these two compounds.
- SFT and Risc-Human predict identical doses for all contaminants of concern, as would be expected from the results of the GTS and the sensitivity analysis on the exposure data.
- The dose predicted by RBCA is approximately half that predicted by SFT and Risc-Human, the main reason for which is that the default commercial worker exposure frequency is approximately half that that modelled in SFT and Risc-Human.
- Vlier-Humaan predicts a lower dose than Risc-Human due to a lower default ingestion rate and a more complex algorithm for modelling the exposure frequency.
- RISC produces a lower dose than the other systems due to the variation in exposure data. The RISC system is unique in providing two sets of default exposure parameters, based on a “typical” exposure and a “reasonable maximum exposure”. For this test site, the “typical” values were selected. Selecting the “reasonable maximum exposure” data would result in doses more similar to those of RBCA, as seen in the GTS.

Dermal contact with soil

All of the systems predict doses based on the contaminants being treated as threshold compounds, i.e. a risk level derived from this dose describes the compound as being a threshold compound and the averaging time is equal to the exposure duration in the dose calculation. Differences are therefore described on this basis, though it should be noted that a number of systems also consider them as non-threshold compounds where the dose may be averaged over a lifetime.

- The results for TCE and PCE are identical across all models for the dermal contact with soil pathway. This is expected since the only chemical properties required for predicting these doses (source concentration, and dermal absorption factor) are defined to be equal for these two compounds.
- RBCA predicts the highest doses for all contaminants. The dermal contact pathway was shown to be similar for RBCA and RISC in the GTS. However the default soil adherence factor and chemical specific dermal absorption factors are greater in RBCA than in RISC. In addition, the chosen exposure dataset in RISC provides a lower exposure for a commercial worker than the default exposure data in RBCA.
- Vlier-Humaan and Risc-Human produce lower doses than RBCA, as expected from the GTS results. The dermal contact doses are greater in Vlier-Humaan due to small differences in the chemical properties and differences in the default exposure parameters.

Inhalation of indoor air

All of the systems predict doses based on the contaminants being treated as threshold compounds, i.e. a risk level derived from this dose describes the compound as being a threshold compound and the averaging time is equal to the exposure duration in the dose calculation. Differences are therefore described on this basis, though it should be noted that a number of systems also consider them as non-threshold compounds where the dose may be averaged over a lifetime.

- RBCA does not predict doses for inhalation pathways; therefore no results are available for comparison.
- As expected from the predicted receptor point concentrations and the results of the GTS the dose predicted by UMS is higher than any other system.
- Risc-Human predicts a greater dose than RISC, even though the predicted receptor point concentration is less and the dose in the GTS is less than that of RISC. This is explained by the variation in the default exposure data. As seen on Figure 26 the Risc-Human exposure data results in a greater increase in the dose over the GTS exposure data than the RISC default exposure data. Though it should be noted that the RISC exposure data is based on a “typical” receptor rather than a “reasonable maximum exposure”, which would result in a higher dose.
- Vlier-Humaan predicts lower doses than Risc-Human, as expected from the receptor point concentrations, GTS and the lower exposure duration Vlier-Humaan incorporates for a commercial worker receptor.

14. TEST SITE 5 - FORMER RETAIL PETROL FILLING STATION

14.1. Introduction

This test site comprises two parcels of land. One is a disused former retail petroleum filling station site, now used as a vehicle repair garage. The other is an engineering manufacturing company adjacent to the filling station, which was previously a part of the filling station. The petrol filling station closed for fuel sales in 1995.

In 1995 a series of investigations showed that the both sites were contaminated with petroleum hydrocarbons due to a relatively recent spill of super unleaded petrol, superimposed over other older contamination. Free phase petroleum hydrocarbon was present on the water table in certain monitor wells. Indoor air and soil gas measurements were collected together with more orthodox soil and groundwater data.

This site provided a good dataset to compare performance of those models that evaluate the vapour intrusion to indoor air pathway.

The following risk models were run for this test site:

RBCA • RISC • Risc-Human • Vlier-Humaan

A site layout plan is presented as Figure 35.

14.2. Conceptual Model

Outlined below are the active source-pathway-receptor linkages to be assessed as part of the modelling process.

14.2.1. Sources

The Contaminants of Concern (COCs) are considered to be:

- Benzene (non-threshold)
- Toluene (threshold)
- MTBE (threshold)

14.2.2. Pathways

The surface of the site comprises made ground of concrete or paving overlying hardcore fill to a depth of 0.5m below ground level. Beneath the made ground is a thin layer of clay and silt between 0.5m and 1.0m in thickness. Underlying the silts and clays are lacustrine sands and gravels down to a maximum depth of 6.0m, which are underlain by the London Clay, a stiff clay formation. The clay extends to a predicted depth of over 50m below ground level.

14.2.3. Receptors

The active exposure scenario at this site is considered to be volatilisation from an off site groundwater plume beneath the engineering manufacturing workshop to indoor air. Workers in the engineering workshop are exposed to these contaminants via inhalation of indoor air.

Other exposure pathways are potentially active at this site, but the site data were considered to be particularly suitable for the assessment of the groundwater to indoor air pathway and not for other pathways.

14.2.4. Input Data

The parameter values used to define the conceptual model are presented in Table 25.

14.2.5. Summary

For this site the exposure of commercial workers to benzene, MTBE and toluene, via inhalation of indoor air impacted by vapours originating from a groundwater source, has been evaluated.

14.3. Results and Discussion

The site conceptual model for test site 5 is presented in Table 25. Indoor air concentrations have been measured in the engineering workshop building, with the range of measured concentration presented below:

Contaminant of Concern	Measured Indoor Air Concentration (Minimum)	Measured Indoor Air Concentration (Maximum)
	(mg/m ³)	(mg/m ³)
Benzene	0.0019	0.0049
MTBE	0.0006	0.0017
Toluene	0.0107	0.0236

This site has been used to demonstrate the levels of conservatism incorporated into each system in modelling the transport of contaminants from groundwater beneath a site into an overlying indoor air space. In this case the measured groundwater concentration beneath the engineering workshop is used rather than model derived concentrations.

Receptor point concentrations and doses are presented on Figures 36A and B respectively with the results tabulated in Table 26.

14.3.1. Receptor Point Concentrations

All of the systems produce predictions of receptor point concentrations that are higher than the measured concentrations.

RBCA, RISC (Group I)

RISC and RBCA produce very good estimates, particularly for toluene and benzene, where the predictions are somewhat higher but well within an order of magnitude of the measured values.

Risc-Human, Vlier-Humaan (Group II)

Risc-Human and Vlier-Humaan both predict receptor point concentrations which are greater than the measured indoor air concentrations by one to two orders of magnitude. The higher receptor point concentrations produced by Risc-Human relative to Vlier-Humaan are due to the differences in the F_{oc} and flux calculations as seen in the GTS.

Groups I versus II

The Group II systems predict greater indoor air concentrations in comparison to the Group I systems, predominantly due to the variation in chemical default properties between the groups. In particular, the default Henry's Law Constants are significantly different between models.

RISC, Risc-human and Vlier-Humaan all produce predictions of receptor point concentrations that are in approximately the same ratios of benzene : MTBE : toluene as the measured indoor air values. However, RBCA produces relatively high MTBE concentrations, both relative to the measured concentrations and relative to the other models, due to a variation in Henry's Law Constant with respect to some of the other systems.

14.3.2. Doses

Benzene doses have been presented as both threshold and non-threshold substances, where appropriate.

Doses have not been presented for RBCA, as the indoor air concentration is compared with a unit inhalation risk in the RBCA system, therefore no dose is calculated. A hand-calculation has not been undertaken, as for the GTS, because this requires assumptions about exposure parameters and is not a true representation of the system output.

Differences between predicted doses for each chemical are consistent with receptor point concentration predictions, indicating that the differences in dose are predominantly controlled by the receptor point concentrations. However, the difference between Risc-Human/Vlier-Humaan and RISC is greater at the dose level, due to the difference in default exposure parameters used by each system.

15. CONCLUSIONS AND RECOMMENDATIONS

15.1. Conclusions

The comparison study has evaluated the risk assessment systems widely used across European countries, culminating in the testing of individual models with defined datasets. In order to maximise the understanding of the models, the examination focused on the use of default parameters, the role of individual algorithms to simulate environmental fate and transport, the methodology of calculating doses and the calculations of risk. Comparisons of output were made for Receptor Point Concentrations, Dose Concentrations and Quantified Risk Levels.

Outlined below are the conclusions of the testing carried out for Phase III – Generic Test Site and Phase IV – Case Study Test Sites.

Generic Test Site – General Conclusions

- **When input data were standardised, the models give generally similar Receptor Point Concentrations and Dose Concentrations.**

The generic test site modelling was undertaken by a single group of risk assessors with access to all of the models using a single definitive dataset. This meant that the use of the models was very controlled. In carrying out the project in this way, a number of potential variations between models were identified and standardised where possible (several parameters were hard-wired). The overall result and conclusion was that where this standardisation is undertaken, very similar results could be obtained. This highlights the importance of understanding the input parameters in terms of what role they play in the calculation.

- **The methods for assessment of risk for non-threshold (e.g. carcinogenic) contaminants are different between the models.**

The differences between the models, when dealing with non-threshold contaminants, were seen in two areas of the calculation process. The first was in the definition of “averaging time”, for example some models produce a lifetime average dose whereas other models define the averaging time as equal to the exposure duration. The second area was in the definition of the health criteria, for example some systems use cancer slope factors whereas other systems use tolerable/index doses (comparable to the tolerable daily intake approach used for threshold contaminants). It is important to note that the methodology for the assessment of non-threshold contaminants is likely to be rooted in a national approach.

- **The models do not produce Quantified Risk Level output in a comparable format.**

There are two important facets to explaining this conclusion:

- a. This study identified that some of the risk assessment systems selected were not directly comparable at the Quantified Risk Level calculation stage. This was because of the definition of health criteria. Where the definition of the health criteria was identical, the calculation of the quantified risk levels was very similar. The calculation process from dose to risk level was in all cases relatively simple and the values of risk would have been highly dependent on the calculated dose. Therefore, a comparison of risk values has not been carried out.

- b. The method of presenting the final quantified risk level varies significantly between the systems. For example, some systems produce a risk level that combines one or more exposure pathways, whereas some systems present individual risks for individual exposure pathways. This made comparisons difficult without manipulation of the models or hand-calculation.

Generic Test Site – Pathway Specific Conclusions

- **Soil ingestion algorithms are all similar, in many cases identical. The predicted soil ingestion doses are correspondingly similar.**
This exposure pathway is a relatively simple model, relying on the specification of a soil contaminant concentration and the amount of soil consumed.
- **Vegetable ingestion doses are relatively uniform, with results generally within one order of magnitude.**
The similarity between the results was relatively surprising given the complexity of the modelling process for this pathway. All models are based on a two stage process: i) the uptake of the contaminant into the vegetable; and ii) the quantity of vegetables consumed by the receptor. The models are all broadly based on the same research paper.
- **Dermal contact doses are more variable, being distributed over two orders of magnitude.**
Three different modelling approaches exist when modelling dermal contact pathways. The first method relies on a defined fraction of contaminant absorbed through the skin. The second method relies on the definition of an absorption rate, thereby requiring a specified length of time over which dermal contact occurs. The third method also uses an absorption rate and includes a mass balance calculation.
- **Indoor air models produce the greatest variation, with predicted doses varying over 3 orders of magnitude.**
The number of processes that can potentially affect the migration of vapours into an indoor airspace are significant. Consequently, the models developed to simulate this pathway are diverse and vary in terms of complexity. The simplest model relies on a basic transfer factor to determine the amount of soil gas that becomes indoor air. The most complex models simulate migration of vapours through either pores or cracks in foundation materials. The potential complexity of this exposure scenario has led to the development of some models that reflect this potential complexity and some that adopt more generic assumptions. It is important to note, that when modelling this pathway, understanding of the assumptions and algorithms, and the potential limitations of the model are essential.
- **Groundwater migration models produced generally similar results**
The risk assessment systems that include groundwater migration and contaminant fate and transport algorithms all produced generally similar results. Four of the systems use essentially similar equations. Two of the systems adopted a more simplistic approach to account for attenuation along a pathway based primarily on dilution.

Generic Test Site – Sensitivity Analysis Conclusions

The sensitivity testing comprised the comparison of the results of modelling using the generic test site data set to the results of datasets derived from the system default data sets. This was done to compare the effects of chemical properties, toxicological properties and exposure parameters.

Running the models with the default internal parameter values produces significant findings. Large variations in output were seen when:

- GTS results are compared with the models run with default exposure factors
- GTS results are compared with the models run with default chemical data

Comparison between the effects of certain models was not possible as some of the parameters used in the generic test site included defaults from selected systems and therefore the results were identical. However, the overall conclusion of this process is that the default datasets vary significantly and consequently the results of the modelling processes varied significantly.

It should be noted that the default values specified in some of the risk assessment systems are national standards or guidelines and therefore would not be changed when undertaking risk assessment in that country.

Test Sites

- Use of model default exposure parameters leads to large differences in doses even for those pathways, such as soil ingestion, that produced similar results in the GTS. This demonstrates that the reliance on default data where parameters are unknown, can produce a varied response, thereby potentially producing highly variable conclusions in a site's risk management.
- A greater variation in results was observed in the Test Site modelling than in the Generic Test Site modelling. Some of the differences were attributable to a wider variation between default chemical property databases. This seemed to be particularly the case for the less common contaminants used in the test sites e.g. lindane.
- At Test Site 1, down gradient groundwater concentrations had been monitored. As part of the assessment of this site, comparison was made between the predicted groundwater concentrations and the measured groundwater concentrations. It was concluded that in this single trial, the groundwater fate and transport predictions produced similar concentrations between the models, and were in good agreement with available site data.
- At Test Site 5, indoor ambient air monitoring had been carried out as part of the original site assessment. As part of the assessment of this site, comparison was made between the predicted indoor air concentrations and the measured indoor air concentrations. It was concluded that indoor air models based on the Johnson + Ettinger algorithms predicted concentrations that gave a reasonable match to available site BTEX data for one test site.

15.2. Recommendations

The Sponsors should consider whether they could contribute to the advancement of risk assessment practice in Europe by following up this study with further evaluation of European risk models and/or development of risk assessment methodologies

16. REFERENCES

American Society for Testing and Materials (ASTM), 1995. Standard Guide for Risk-Based Corrective Action Applied at Petroleum Release Sites (ASTM E 1739-95).

Arge Fresenius-focon, 1999. UMS, Exposure to Harmful Substances; Umwelt und Mensch mit Schadstoffen. Chemlog GbR, IfUA GmbH.

Arge Fresenius-focon, 1999. Anhang1: Expositionsszenariospezifische Beschreibung der Ableitungsvoraussetzungen der Transfer und Expositionsdaten. Chemlog GbR, IfUA GmbH.

Briggs, G.G., Bromilow, R.H., Evans, A.A., 1982. Relationships between lipophilicity and root uptake and translocation of non-ionised chemicals by barley. *Journal of Pesticide Science* 13:495-504.

Danish Environmental Protection Agency, 2002. Environmental Guidelines No. 7. Guidelines on Remediation of Contaminated Sites. Danish Ministry of the Environment.

Environment Agency, 1999. Methodology for the Derivation of Remedial Targets for Soil and Groundwater to Protect Water Resources. (Marsland, P.A., and Carey, M.A.) Environment Agency R&D Publication P20.

Environment Agency, 2002. CLEA 2002 version 1.3. Department of Environment, Food and Rural Affairs and Scottish Environment Protection Agency.

Environment Agency, (in press). Methodology for Comparison of CLEA with other human health risk assessment packages. Environment Agency NGWCLC Report NC/06/07. LQM Press.

Evans, D., Hers, I., Wolters, R.M., Boddington, R.T.B., Hall, D.H., 2002. Vapour Transport of Soil Contaminants. Environment Agency and Institute of Petroleum Research and Development Technical Report P5-018/TR.

Ferguson, C., Darmendrail, D., Freier, K., Jensein, B.K., Jensen, J., Kasamas, H., Urzelai, A., and Vegter, J. (editors), 1998. Risk Assessment for Contaminated Sites in Europe. Volume 1: Scientific Basis. LQM Press, Nottingham.

Ferguson, C.C., and Kasamas, H. (editors), 1999. Risk Assessment for Contaminated Sites in Europe. Volume 2: Policy Frameworks. LQM Press, Nottingham.

Golder Associates. ConSim 1.05 Manual. Environment Agency.

Howard, P.H., Boethling, R.S., Jarvis, W.F. Meylan, W.M., Michalenko, E.M., 1991. Handbook of Environmental Degradation Rates, Lewis Publishers Inc. Chelsea. MI.

Marsland, P.A., Carey, M.A., 1999. Methodology for the Derivation of Remedial Targets for Soil and Groundwater to Protect Water Resources. Environment Agency R&D Publication P20.

Spence L & BP Oil, 2001. Risk-Integrated Software for Clean-ups, RISC User's Manual - Version 4.0.

Swartjes, F.A., 2002. Human exposure model comparison study: State of play. Presentation Paper

URS Dames & Moore, 2002. Giuditta version 2.0 User's Handbook. Province of Milan.

Vik, E.A. and Bredeveld, G., 1999. Report 99:06, Guidelines for the Risk Assessment of Contaminated Sites. SFT (ISBN 82-7655-192-0).

VROM, 2000. Circular on Target Values and Intervention Values for Soil Remediation. (Dutch) Ministry for Housing, Spatial Planning and Environment.

Whittaker, J.J., Buss, S.R., Herbert, A.W. & Fermor, M., 2001. Benchmarking and Guidance on the Comparison of Selected Groundwater Risk Assessment Models. Environment Agency NGWCLC Report NC/00/14.

Zalaski, R., Gephart, L., 1999. NICOLE Exposure Factors Sourcebook for European Populations, with focus on UK Data.

Rikken, M.G.J., Lijzen, J.P.A., Cornelese, A.A., 2001. Evaluation of Model Concepts on Human Exposure. (Netherlands) National Institute of Public Health and the Environment (RIVM-Report 711701 022).