Challenges to Collecting Health Data for Secondary Use: A Case Study

---

Patricia Heckmann

A dissertation submitted to the University of Dublin,

in partial fulfilment of the requirements for the degree of

Masters of Science in Health Informatics.

2013

Declaration

I declare that the work described in this dissertation is, except where otherwise stated, entirely my own work, and has not been submitted as an exercise for a degree at this or any other university.

Signed: ______________________________ Date: ___________

Patricia Heckmann

Permission to lend and/or to copy

I agree that the Trinity College Library may lend or copy this dissertation upon request.

Signed: ______________________________ Date: ___________

Patricia Heckmann

Summary

"Change is not made without inconvenience, even from worse to better."

Richard Hooker (1554 – 1600)

There is a growing recognition of the need to base healthcare policies on accurate, detailed and timely data. These data are predominantly generated through clinical documentation in the process of direct patient care. This results in an increasing demand on healthcare organisations involved in direct patient care to collect and report these data for secondary use.

The literature review revealed that good quality data is essential to manage public health programmes and that there are many widely recognised challenges to the secondary use of data. This research utilised an embedded single-case study, to focus on describing the context for collection of health data for secondary use in Ireland, on the actual and perceived barriers to the collection of these data, and on opportunities which would facilitate that collection. The research method and data collection tools utilised included interviews, questionnaires and time studies, in addition to the literature review.

It was concluded that Ireland, similar to other countries, is experiencing several challenges in the collection of health data for secondary use. The sociocultural challenges to implementing health data collections for secondary use are as daunting as the technical and economic ones. An alternative strategy for data collection will be required to ensure a sustainable process going forward. Collection of health data must be integrated into the documentation of the clinical care pathway of the patient if it is to remain sustainable and accommodate the increasing requirements of a well-managed health service. This will require the implementation of the essential building blocks of a national health infrastructure network in addition to health data legislation, a data stewardship model, and best practice in national health data collection and data sharing. A comprehensive and well-managed change management process, in addition to accountable leadership both at local and national level, is necessary to ensure that systems implemented meet their full potential.

Acknowledgements

I would like to sincerely thank everyone who supported and assisted me during this research assignment, particularly:

- My husband Henrik, and children Jennifer, Caoimhe and Maitiú, without whose support, understanding and encouragement I could not have completed this assignment.
- My mother Eileen for being so understanding of my absences and lack of communication.
- The memory of my late father Cornelius who lived by the principles of ongoing and further education, both formal and informal.
- My current and former managers, Dr. Susan O'Reilly and Prof. Ciaran Meegan who were both very supportive of my studies.
- My current and former work colleagues who took time out of their busy schedules to help me with this research, and to give me invaluable feedback.
- My research supervisor Dr Lucy Hederman for her guidance, help and support throughout this research assignment, and course of study.

To all of you, many thanks!

> "*To reach a port we must set sail – Sail, not tie at anchor. Sail, not drift.*"
>
> *Franklin D. Roosevelt (1882 – 1945)*

Table of Contents

List of Tables 1

List of Figures 3

List of Boxes 4

Abbreviations 5

Chapter 1. Background 6
- *1.1 Introduction 6*
- *1.2 Research Aims 7*
- *1.3 Motivation for the Research 8*
- *1.4 Overview of Dissertation 8*

Chapter 2. Literature Review 10
- *2.1 Introduction 10*
- *2.2 Secondary Use of Health Data 11*
- *2.3 Cancer Control and Treatment Cost 12*
- *2.4 Public Health, Cancer Control and the Requirement for Data 13*
- *2.5 Location of Health Data 14*
- *2.6 The Role of Information Technology in Health Data Reporting 14*
- *2.7 Data Intelligence 17*
- *2.8 Challenges to the Secondary Use of Health Data 18*
- *2.9 Identification of Stakeholders 27*
- *2.10 Key Findings of the Literature Review 27*
- *2.11 Conclusion 28*

Chapter 3. The Selected Case 29
- *3.1 Background 29*
- *3.2 The Selected Case 29*
- *3.3 Conclusion 33*

Chapter 4. Research Methodology 34
- *4.1 Introduction 34*
- *4.2 Role of Literature Review in Research 34*
- *4.3 Choice of Research Design and Strategy 34*

4.4 Identification of Stakeholders ... 38
4.5 Research Methods ... 39
4.6 Conclusion ... 43

Chapter 5. Quantitative Research Findings ... 45
5.1 Introduction ... 45
5.2 Questionnaire Results ... 45
5.3 Time Study Results ... 54
5.4 Personnel Resources Required for Data Input ... 56
5.5 System Reporting Ability ... 59
5.6 Data Fields of the Selected Case ... 60
5.7 Key Quantitative Findings ... 60
5.8 Conclusion ... 62

Chapter 6. Qualitative Research Findings & Discussion ... 63
6.1 Introduction ... 63
6.2 Semi-structured Interviews - Data Analysis ... 63
6.3 Challenge - Inefficiency ... 66
6.4 Challenge - Variations in Measurement Systems ... 69
6.5 Challenge - Organisational and Cultural Issues ... 70
6.6 Challenge - Technological Barriers for Electronic Health Records ... 74
6.7 Challenge - Economic Pressures and Competing Priorities ... 75
6.8 Data Ownership, Data Protection and Identifiers ... 78
6.9 Conclusion ... 80

Chapter 7. Conclusion ... 81
7.1 Research Summary ... 81
7.2 Implications of the Research Findings ... 81
7.3 Limitations of the Research Findings ... 81
7.4 Research Conclusions and Recommendations ... 82
7.5 Recommendations for Further Research ... 84

References ... 85

Appendix 1. Ethics approval ... 98

Appendix 2. Gatekeeper approval ... 99

Appendix 3. Informed consent form and information sheets 100

Appendix 4. Minimum data set – PCRS reimbursement portal.............................. 109

Appendix 5. Data field comparison 110

Appendix 6. County code data dictionaries 112

Appendix 7. Data collection tool - time study 114

Appendix 8. Coded thematic analysis... 118

Appendix 9. Hospital pharmacist questionnaire screen shots................................. 137

Appendix 10. Systems vendor questionnaire screen shots 143

Appendix 11. Location of data in hospital systems 151

Appendix 12. Hospital and NCCP data responses.. 152

Appendix 13. Hospitals included in the research.. 153

Appendix 14. Proposed interview topics 154

Appendix 15. Semi-structured interviews – notes record form 155

Appendix 16. Stakeholder details 162

List of Tables

Table 2.1: Literature search 10

Table 2.2: Components of a national framework for secondary use of health data 12

Table 2.3: Primary and secondary uses of an electronic health record system 16

Table 2.4: AMIA’s proposed principles for clinical data capture and documentation... 17

Table 2.5: AHRQ data collection challenges........ 18

Table 2.6: Problems caused by disparate data element definitions and taxonomies 19

Table 2.7: Comparison of health identifiers in Australia and Ireland 23

Table 2.8: Data stewardship principles........ 26

Table 2.9: The three project interests as defined by Prince2TM 27

Table 3.1: SACT reimbursement scheme data entry steps 30

Table 4.1: Selection of the case 37

Table 4.2: The three project interests........ 38

Table 4.3: Research objectives of the semi-structured interviews........ 39

Table 4.4: Research objectives of the questionnaires 41

Table 4.5: Research objectives of the time study 41

Table 4.6: Components of the time study and the research approach 42

Table 4.7: Work elements of the data input process........ 43

Table 5.1: Hospital questionnaire sections aligned to research objectives........ 46

Table 5.2: Classification of response rates to the four main sections........ 46

Table 5.3: Selected free text comments on the difficulties of data extraction........ 53

Table 5.4: Estimated annual national hospital activity data 54

Table 5.5: Data input work elements 55

Table 5.6: Time study results........ 56

Table 5.7: Web portal work processes........ 56

Table 5.8: Assumptions for personnel resource calculations 57

Table 5.9: Scenarios utilised to estimate the data input time per annum........................ 57

Table 5.10: Estimated time required per annum for data input 58

Table 5.11: Vendors’ responses to report export format 59

Table 5.12: Vendors’ responses to available reports 60

Table 6.1: Categorisation of interview topics 65

Table 6.2: Interviewees’ views on the reuse of existing electronic data 68

List of Figures

Figure 3.1: Web portal patient search screen 31

Figure 3.2: Web portal patient details screen 31

Figure 3.3: Web portal patient review screen 32

Figure 3.4: Web portal clinical details registration screen 32

Figure 3.5: Web portal claim screen 33

Figure 5.1: Hospital questionnaire participation 45

Figure 5.2: Hospital questionnaire response rate to the four main sections of the hospital questionnaire 47

Figure 5.3: ICT systems in use in hospitals to support SACT services 47

Figure 5.4: Stored locations of demographic data 48

Figure 5.5: Stored locations of clinical and drug data 49

Figure 5.6: Hospital SACT preparation systems in use 50

Figure 5.7: Data conflicts - demographic data 50

Figure 5.8: Data conflicts - clinical and drug data 51

Figure 5.9: Hospital activity responses graded as complete, partial or no response 51

Figure 5.10: Hospital activity data by source - NCCP or hospital questionnaire 52

Figure 5.11: Participants' experience in patient registration 55

Figure 6.1: Stakeholders interviewed 63

List of Boxes

Box 6.1: Interview quotations on inefficiency and duplication of work 66

Box 6.2: Interview quotations on the role of ICT in data collection 67

Box 6.3: Interview quotations on data quality 68

Box 6.4: Interview quotations on the subject of mandatory reporting 69

Box 6.5: Interview quotations on variations in measurement systems........................... 70

Box 6.6: Interview quotations on the requirement for data 70

Box 6.7: Interview quotations on change management and incentives 72

Box 6.8: Interview quotations on education and training of staff 72

Box 6.9: Interview quotations on challenges to full utilisation of software systems 73

Box 6.10: Interview quotations on procurement approaches 73

Box 6.11: Interview quotations on challenges to ICT implementation 74

Box 6.12: Interview quotations on standards and interoperability 75

Box 6.13: Interview quotations on technology strategies.. 75

Box 6.14: Interview quotations on competing priorities 76

Box 6.15: Interview quotations on cost of data collection 77

Box 6.16: Interview quotations on minimum data set requirements 77

Box 6.17: Interview quotations on business intelligence 78

Box 6.18: Interview quotations on data protection... 79

Box 6.19: Interview quotations on unique identifiers... 80

Box 6.20: Interview quotations on data ownership 80

Abbreviations

AHRQ	Agency for Healthcare Research and Quality
AMIA	American Medical Informatics Association
BCCA	British Columbia Cancer Agency
CMOD	Centre for Management and Organisational Development
DoH	Department of Health
DPC	Data Protection Commissioner
eHealth	Electronic Health
EHR	Electronic Healthcare Record
ESRI	The Economic and Social Research Institute
HIE	Health Information Exchange
HiMSS	Healthcare Information and Management Systems
HIPE	Hospital In-Patient Enquiry Scheme
HIS	Hospital Information System
HIQA	Health Information and Quality Authority
HSE	Health Services Executive
ICT	Information and Communications Technology
IOM	Institute of Medicine
KPI	Key Performance Indicator
MDM	Multidisciplinary Meeting
MRN	Medical Record Number
NCCP	National Cancer Control Programme
NCIS	National Cancer Information System
NCRI	National Cancer Registry, Ireland
NEHTA	National eHealth Transition Authority
NHIN	National Health Infrastructure Network
NHS	National Health Service United Kingdom
NTPF	National Treatment Purchase Fund
PCRS	Primary Care Reimbursement Service
PHI	Personal Health Information
PPSN	Personal Public Service Number
SACT	Systemic Anti-Cancer Therapy
SNOMED CT	Systematized Nomenclature Of Medicine Clinical Terms

Chapter 1. Background

1.1 Introduction

Healthcare is a complex data-intensive environment generating terabytes of data to support patient care and treatment (Safran et al., 2007, HIQA, 2011b, Barton et al., 2011). The aspiration to achieve a “learning healthcare system” is creating an increasing requirement for data from various initiatives and organisations, both local and national (AHRQ, 2006, HIQA, 2013c, HSE, 2013c), to ensure quality, safety and value in healthcare (IOM, 2007).

Health data are primarily generated through clinical documentation in the process of direct patient care (Safran et al., 2007, Cimino, 2011). These data can then be made available for “secondary use”, in non-clinical applications such as disease surveillance, quality and performance measurement and reimbursement, amongst others, which often differ from the original collection intent, (Safran et al., 2007, Barton et al., 2011, Health Industries, 2009, IOM, 2003, AHRQ, 2006). The aggregation of these data allows the coordination and management of the health services and informs strategic decision-making (Wilkinson et al., 2007, Department of Health, 1998, Bloomrosen and Detmer, 2008, Turner, 2009) and is seen as a means to improve the quality, safety, and efficiency of healthcare (Barton et al., 2011, ONC, 2010, Stroetmann et al., 2006).

Despite this, access to and use of health data poses complex ethical, political, technical, and economic challenges (Safran et al., 2007). “Data stewardship” is an essential component of a national framework for data reuse as it balances the rights of an individual to have their personal health information (PHI) protected and their wish to have an improved health service (Bloomrosen and Detmer, 2008). The need for a national health infrastructure network utilising data standards, is fundamental to allow the exchange of patient information across healthcare settings and to meet healthcare needs (IOM, 2003, IOM, 2001b, HIQA, 2011b) and to facilitate the “collect once, use many times” paradigm of secondary use (Barton et al., 2011, HIQA, 2013c). This places data standards and interoperability at the centre of any discussion on health data reuse, with both elements being key to the elimination of information silos generated by legacy and non-integrated health care systems (Cimino, 2011, HIQA, 2013c, Barton et al., 2011). The collection and reporting of data presents challenges, including staffing resources, and can place economic and managerial pressures on hospitals through duplication of effort, increased expense and lost opportunities (AHRQ, 2006, Wilkinson et al., 2007, Bradley et al., 2006). These can be reduced through the use of electronic healthcare records (EHRs) (AHRQ, 2010, Lynn, 2013, McLean, 2006, Cusack et al., 2012, Deshmukh et al., 2011).

Ireland, similar to other countries, has increasing requirements for health data driven by the requirements of the Health Services Executive (HSE) and the Department of Health (DoH), amongst others (Balanda et al., 2008). The requirement for an Irish national health infrastructure network, and an integrated approach to Irish health data with a view to removing the challenges to the secondary use of health data, has been acknowledged as necessary to meet increasing demand for data (HIQA, 2012, HIQA, 2011, Balanda et al., 2008, DoH&C, 2008, DOH&C, 2004, Comber, 2006, Staines et al., 2001, HSE Consultants in Public Health Medicine Clinical Programmes Group, 2011).

Mindful of the increasing need and demand for health data, and the challenges to its collection, the researcher considered that information on the contemporary challenges to the collection of health data for secondary use was fundamental in identifying opportunities to improve the data collection processes, so as to ensure that these processes will meet the growing needs of the management of health services. The primary research question chosen for the purpose of the study was:

What are the challenges and facilitators to the collection of health data for secondary use in Ireland?

The case utilised to explore these challenges allowed an opportunity to further analyse one subunit of the challenges, and attempted to answer the following question:

What personnel resources are required to input the data of the selected case?

1.2 Research Aims

1. To undertake a comprehensive review and critical appraisal of the available literature.
2. To investigate the selected case in the areas of:
    - The challenges and facilitators to the collection of health data for secondary use.
    - The personnel resources required to input the data of the selected case.

3. To reflect on the findings, to draw conclusions in relation to the challenges to health data collection for secondary use, and to make recommendations beyond the study domain.

1.3 Motivation for the Research

The primary motivation for this research topic came from the author’s experience of the challenges involved in data collection processes and her interest in the role of information and communications technology (ICT) in improving these processes.

The preliminary literature review revealed there are many challenges to the collection of health data for secondary use, making it a complex phenomenon suitable for research in a real-life context. It also showed that there are many Irish government reports and recommendations, based on international literature, describing and recommending the changes needed to improve the collection of health data for secondary use in Ireland; but little published research on the contemporary Irish challenges to the collection of health data for secondary use. This research will attempt to address that deficit by identifying the contemporary challenges, as experienced by those required to collect health data for secondary use. This research will provide evidence of the challenges to the secondary use of health data and inform prioritisation of the steps and projects essential to improve this phenomenon.

1.4 Overview of Dissertation

Chapter 2: Literature Review: This chapter reviews relevant and scholarly literature relating to the research area. The main themes explored focussed on the collection of health data for secondary use in the domain of cancer control and public health.

Chapter 3: The Selected Case: This chapter describes the selected case, the SACT reimbursement scheme, some relevant changes in the Irish health service, the drivers in the introduction of the selected case and the process of data input. The selected case, set in the domain of cancer care, requires the collection and reporting of data for reimbursement; a secondary use of data. These data originate in the secondary care institutions involved in the provision of cancer services and are reported to the Primary Care Reimbursement Service (PCRS).

Chapter 4: Research Methodology: This chapter outlines the approach taken to answering the research questions, the research methodology adopted, the reasons for its selection and its limitations. An embedded single-case study design was used to illustrate the contemporary challenges to, and potential facilitators of, the collection of health data for secondary use. The embedded subunit research further explored the personnel resources required to input the data of the selected case. The research method and data collection tools utilised included interviews, questionnaires and time studies, in addition to the literature review.

Chapter 5: Quantitative Research Findings: This chapter presents the research findings from the questionnaire responses and the outcome of the time study, and integrates these with the literature review. This allowed the researcher to explore the personnel resources required to input the data. It also ascertained the format in which the required data was stored, in addition to demonstrating the range of computer systems in use to support cancer services in the hospitals.

Chapter 6: Qualitative Research Findings & Discussion: This chapter presents the research findings from the semi-structured interviews and integrates these with the literature review and the data from the quantitative research. Integration of these findings allowed the researcher to identify the challenges to the collection of health data, with a view to informing future data collections and approaches to improving the collection of health data for secondary use.

Chapter 7: Conclusion: This chapter presents the conclusions and recommendations of this dissertation, discusses the research limitations and identifies areas where further research may be needed.

Chapter 2. Literature Review

2.1 Introduction

As stated in Section 1.4, this chapter will explore the relevant and scholarly literature which describes the research problem. A literature review allows the researcher to gather information about research design, data collection and analysis methods, as well as assembling data and conclusions across research (Fink, 1993, Garrard, 2010, Fink, 2009) and determining the scope of the inquiry (Creswell, 2009). The literature review also helps explain any research findings, in addition to identifying current knowledge which justifies the requirement for additional research (Fink, 2009, Creswell, 2009).

The researcher, acknowledging the importance of a literature review, undertook an extensive review in addition to interviewing experts and stakeholders in the area in order to answer the following questions:

What are the challenges and facilitators to the collection of health data for secondary use in Ireland?

What personnel resources are required to input the data of the selected case?

2.1.1 Search strategy

A comprehensive review and critical appraisal of the available literature was undertaken in the areas of secondary use of health data, focussed on the challenges associated with the collection of health data for secondary use, the requirement for health data in the domain of cancer control and public health and the impact of data collection on personnel resources. The databases, journals and grey literature sources that were searched are detailed in Table 2.1 below.

Table 2.1: Literature search

Databases	Journals*	Institutions/Societies
PubMed	International Journal of Quality in Health Care*	NCCP
Google Scholar		HSE
Embase		HIQA
Biomed central		AHRQ
CINAHL		NHS
Science Direct		BCCA & Ontario Cancer Programmes
		DoH&C
		Australia – NEHTA
*Journals not indexed in the databases selected.		

The main search terms used to discover literature of relevance to the subject in the initial searches included “healthcare”, “data”, “secondary use”, “challenges”, “stewardship”, “standards”, “cancer control”, “public policy”, “framework”, “registry”, “aggregation”, “collection”, “reporting”, “stakeholders”, “facilitators”. All searches were limited to human subjects, the English language and the timeline 1995-2012 to gather relevant and recent data, using combinations of Boolean Operators. The literature search was refreshed at intervals to include emerging information.

The reference abstracts from the initial search were reviewed and those most closely related to the research subject selected. The reference lists of these key articles were reviewed to discover other relevant articles repeatedly cited by the scholars. These articles were then retrieved and reviewed.

The researcher found there was an abundance of literature in relation to secondary use of data, especially in the area of clinical trials, clinical research and public health, but few articles focussed on the real-world activity of health data aggregation and reporting, or the contemporary challenges in Ireland.

2.2 Secondary Use of Health Data

In 2007, the American Medical Informatics Association (AMIA) defined the secondary use of health data as “non-direct care use of PHI including, but not limited to, analysis, research, quality/safety measurement, public health, payment, provider certification or accreditation, and marketing and other business including strictly commercial activities” (Safran et al., 2007).

The potential to reuse data gathered to support patient care and treatment for secondary uses in non-clinical applications has the potential to transform healthcare through improvements in quality, safety and efficiency (Barton et al., 2011, Stroetmann et al., 2006, ONC, 2010, HIQA, 2012c, Jones et al., 2012, Safran et al., 2007, AHRQ, 2006, IOM, 2003, IOM, 2001a).

The AMIA recommends that secondary use of health data and a national framework on the uses of health data are prioritised by policy makers (Bloomrosen and Detmer, 2008, Cusack et al., 2012). This situation is supported in Ireland, where the Health Information and Quality Authority (HIQA) recently conducted an international review of the secondary use of PHI in the context of information governance (HIQA, 2012c, HIQA, 2012b). In addition, the DoH is in the process of enacting the Health Information

Bill which is expected to contain a legislative basis for unique health identifiers, in addition to a framework for the use of health data (DOH, 2013b).

In 2007 the AMIA moved beyond the concept of data ownership to “data stewardship” which places the emphasis on data security, access and control (Bloomrosen and Detmer, 2008). This is outlined in the proposal for a national framework for secondary use, as detailed in Table 2.2 below.

Table 2.2: Components of a national framework for secondary use of health data (Safran et al., 2007)

- Transparent policies and practices for the secondary use of health data
- Focus on data control, rather than data ownership per se
- Consensus on privacy, policy, and security
- Public awareness
- Comprehensive scope (beginning with a taxonomy)
- National leadership

2.3 Cancer Control and Treatment Cost

The term cancer is used to describe a group of over a hundred diseases that occur when malignant forms of abnormal cell growth develop in one or more body organs (National Cancer Forum, 2006). Many countries have implemented population-based programmes, termed cancer control, as a component of public health strategies, to reduce the incidence and mortality from cancer in addition to minimising the impact on those who develop cancer (Elwood and Sutcliffe, 2010). Aspects critical to the success of a cancer control strategy include management systems that facilitate information gathering, monitoring and evaluation (Caron, 2010, HIQA, 2012d).

2.3.1 Cost of cancer treatment

The increasing cost of funding new and often expensive cancer drugs, termed systemic anti-cancer therapy (SACT), poses challenges for cancer policy and cancer services, and generates tension between the interests of an individual and the population as a whole (Browman et al., 2010). In some countries SACT funding lies with the provider institution, leading to postal code treatment decisions and inequitable patient access (Richards, 2010). Other countries have taken a centralised approach to SACT funding, recognising that effective systems which monitor SACT utilisation are essential to protect drug budgets, while ensuring equity of treatment for the entire population. These systems, in addition to contributing to good management of scarce financial resources, also allow for the introduction of evidence-based guidelines which improve both patient access and outcomes (O'Reilly and Venkatesh, 2010, O'Reilly, 2013).

2.3.2 Cancer control in Ireland

Cancer, its prevention, diagnosis and treatment are a major challenge for Irish society. Each year approximately 20,000 Irish people develop cancer and 7,500 die from the disease. The number of cases detected per annum is expected to increase from under 14,000 in 2000, to over 28,000 in 2020 (National Cancer Forum, 2006, Comber, 2006b). The Irish National Cancer Control Programme (NCCP) was established in 2007 to transform the delivery of cancer care, and to ensure that cancer services meet the highest standards based on successful international models (HSE, 2012e, Coleman et al., 2011, WHO, 2013, National Cancer Forum, 2006).

2.4 Public Health, Cancer Control and the Requirement for Data

Public health has been described as "the science and art of preventing disease, prolonging life and promoting health and efficiency through organised community efforts" (Winslow, 1920). The core functions of public health are underpinned by a requirement for health data (HIQA, 2012d, BC Population Health and Wellness, 2005). These data are required, not only at the point of service delivery but also at the point of decision-making, so as to achieve a "learning healthcare system" (IOM, 2003). They must be in a format that facilitates this decision-making so as to drive improvements in safety, quality effectiveness and sustainability in addition to evaluation and audit (Department of Health, 1998, Wilkinson et al., 2007, Bloomrosen and Detmer, 2008, Grossman et al., 2008, Pevnick et al., 2012, Balanda et al., 2008, HIQA, 2012e). This also leads to better-informed patients and a better-informed public through consistent availability and use of health information (National Cancer Forum, 2006, Browman et al., 2010).

Ireland, similar to other countries, has an increasing requirement for health data (HSE, 2013c, HSE, 2013a, DOH, 2004, DOH, 2008, HIQA, 2013c). The implementation of the HSE's clinical care programmes and the broadening data requirements of the cancer control programme are some of the factors driving and increasing the requirement for data (O'Reilly, 2013).

2.4.1 Irish cancer control data requirements

Cancer control, as described in section 2.3.2, is one of Ireland's public health programmes that relies heavily on data collected in direct patient care so as to fulfil its public health obligations (Coleman et al., 2011, O'Reilly, 2013). The NCCP's increasing requirement for health data is evident from the introduction of new key performance indicator (KPI) data collections for prostate and lung cancer in 2012 (HSE, 2012c), in addition to the introduction of the SACT reimbursement scheme, as discussed in Chapter 3, which requires hospitals to submit data to ensure their funding.

The National Cancer Registry (NCRI) has the benefit of being comprehensive, accurate and population-based, but is retrospective, and historically produces its reports a year or more after the year of diagnosis, which does not fulfil the NCCP's requirement for timely data as discussed in Section 2.4 . It has the benefit of being exempt from the provisions of the Health (Provision of Information) Act 1977 and thus can collect the patient-specific cancer health data without the need for patient consent (Data Protection Commissioner, 2013b).

There is a recognition that a richer data set, in the form of a national cancer information system (NCIS), incorporating a core national cancer dataset, is required to meet the NCCP and the NCRI data requirements. This information system could operate independently, or as a subset, of a national health infrastructure network (NCIN, 2013, McDevitt and Comber, 2009). NCISs may contain a large variety of patient-level data, a subset of which relates to systemic anti-cancer therapy (SACT) (NHS CIU, 2011, McDevitt and Comber, 2009). These data are frequently extracted from electronic systems in use in the provider institutions to improve the safety of SACT preparation (SHPA, 2005, ISMP, 2013, ASHP, 2008) and prescribing (NCAG, 2009, Carrington et al., 2010).

2.5 Location of Health Data

Health data are primarily generated through clinical documentation in the process of direct patient care (Safran et al., 2007, Cimino, 2011), allowing the potential for reuse, as described in section 2.2. This documentation may be held in paper charts, electronic health records (EHRs), or a mixture of both. Its extraction, manually or electronically, is necessary to meet the data reporting requirements of the institution, in addition to secondary users of these data (AHRQ, 2006, HIQA, 2011a, HIQA, 2012c).

2.6 The Role of Information Technology in Health Data Reporting

The need for a national health infrastructure network (NHIN) utilising data standards, is recognised as fundamental in order to allow the exchange of patient information across healthcare settings and to meet healthcare needs (IOM, 2001b, IOM, 2003, HIQA, 2011b, Shapiro, 2007). Elements of such a network include electronic health record

(EHR) systems, as discussed in Section 2.6.1, in addition to data standards, as discussed in Section 2.8.4.2 (IOM, 2004). The “collect once, use many times” paradigm of data reuse is advocated as a way towards a national health information strategy (Barton et al., 2011).

The introduction of national standardisation, in the form of NHINs, has been presented as extracting more value from previous IT investment (Health Industries, 2009) in addition to having the potential to generate net savings, estimated to range from $21.6 – $77.8 billion per year in the United States, dependent on the levels of implemented health information exchange (HIE) (Walker et al., 2005). HIEs promise cost and quality improvements, yet there is a deficit in substantial and consistent empirical demonstrations of their effectiveness and few examples exist of this being achieved as a by-product of clinical care (AHRQ, 2006, Vest and Gamm, 2010, Kellermann and Jones, 2013).

The deficit of functional HIE networks results in institutions using proprietary data feeds directly from individual hospitals and other data providers, instead of data directed to them through a clinical HIE implementation (Shapiro, 2007). Equally the absence of agreed standards and terminologies, as discussed in Section 2.8.4.2, have led to projects that map institutionally “siloed” data, which achieve an interim solution and provide data aggregation for the purposes of research, amongst others (Lowe et al., 2009, Hernandez et al., 2009, Rea et al., 2012). These developments achieve the required outcome for the systems and institutions involved, but are limited in scalability as mapping, which is resource-intensive, and is required to include data from additional systems (AHRQ, 2006, Kush et al., 2007).

2.6.1 The role of electronic health records in health data reporting

Electronic health records (EHRs) can be described as “digitally stored healthcare information about an individual’s lifetime with the purpose of supporting continuity of care, education and research, and ensuring confidentiality at all times” (Iakovidis, 1998). They may be seen as a longitudinal record of all care provided to the patient, but this record is usually held within individual institutions rather than across institutions (McLean, 2006).

EHRs can facilitate a culture of health data management and sharing, which in turn can create a tremendous impact on patient care and health outcomes (Altarum Institute, 2011, AHRQ, 2010, Kellermann and Jones, 2013). They have many primary and secondary uses, similar to paper medical records, as detailed in Table 2.3, yet they allow access to the relevant data in a timely fashion, thereby making it easier to realise those uses (Lynn, 2013). This has resulted in a growing interest in the reuse of EHR data for secondary uses (Deshmukh et al., 2011, EHR4CR, 2011).

Table 2.3: Primary and secondary uses of an electronic health record system (IOM, 2003)

Primary and Secondary Uses of an Electronic Health Record System	
Primary Uses	**Secondary Uses**
Patient Care Delivery	Education
Patient Care Management	Regulation
Patient Care Support Processes	Research
Financial and Other Administrative Processes	Public Health and Homeland Security
Patient Self-Management	Policy Support

The major value of integrated EHRs is that they enable capture of clinical data as part of the overall workflow (McLean, 2006), thereby allowing its reuse (Ayatollahi et al., 2009, 1992, Shekelle et al., 2006). Successful integration of data collection into clinical workflow is essential to minimise workflow disruption and redundant data capture. This can be achieved through the implementation of user-friendly systems (Kellermann and Jones, 2013) and the AMIA's principles of clinical data capture and documentation, as detailed in Table 2.4 (Cusack et al., 2012, Porter, 2013). That aside, institutions and people must be open to process change and revision to ensure that implemented systems meet their full potential (Kellermann and Jones, 2013).

Table 2.4: AMIA’s proposed principles for clinical data capture and documentation (Cusack et al., 2012)

Clinical data capture and documentation should:

1. Be clinically pertinent, patient-centric, and represent an individual’s lifetime health and healthcare.
2. Support capture of high-quality information that is accurate, relevant, confidential, reliable, valid, complete, and secure.
3. Be efficient and usable while enhancing the healthcare organization’s and the care team’s overall efficiency, effectiveness and productivity.
4. Support multiple downstream uses as a by-product of the recording of care delivery, including quality measurement, performance improvement, population health care delivery, policymaking, research, education, and reimbursement.
5. Enable joint patient-provider decision-making, team collaboration, care process management, and advanced clinical decision support.
6. Enable collection of data and interpretation of information from multiple sources as appropriate and necessary, including nuanced medical discourse, structured items, and data captured in other systems and devices.
7. Enable automation of data capture and documentation which should be optimized whenever appropriate, allowing human beings to focus on gathering and entering data that cannot be effectively collected by automated tools (e.g. automated acquisition of data from biomedical devices).

Data abstraction from EHRs is a core functionality required to produce both user-defined and standard reports (Gelmon and Droppers, 2008). EHRs have been seen as a way to allow enhanced data collections and access (Agrawal, 2002). The documentation burden should be lowered through the use of data that is structured, computable and semantically interoperable, collected at the point of patient care and reused for secondary purposes (Ottosen, 2012), thus reducing the requirement for clerical staff resources (Sidorov, 2006, Miller et al., 2005). Studies have shown direct financial benefits with the use of EHRs through reduction or avoidance of costs in the areas of labour costs for coding and billing, searching for paper charts and transcription costs (Menachemi and Collum, 2011, Silow-Carroll et al., 2012, Iakovidis, 1998, Wang et al., 2003, Ottosen, 2012, Kellermann and Jones, 2013). Although electronic extraction of this data is preferable to manual extraction, there remain many challenges to its implementation, as detailed in section 2.8.4.2.

While Ireland currently does not operate a nationally agreed EHR, there are disparate EHRs in use as individual hospitals have purchased, or built, their own systems (HIQA, 2011a).

2.7 Data Intelligence

The expanding requirements for health data, data integration and the linkage of datasets are creating vast data sources in many domains, including healthcare (PWC, 2012) and require pertinent, efficient and cost-effective processing for enhanced insight and decision making (McKinsey, 2011). Access to data and the timely dissemination of analysed data in a usable format, including to those institutions providing the data, is essential to ensure that the data value is maximised (HIQA, 2013c, Wilkinson et al., 2007). This enables utilisation of national data at local level (PWC, 2012) and promotes the value of national data collection (Wilkinson and McCarthy, 2007).

2.8 Challenges to the Secondary Use of Health Data

Access to, and use of health data poses complex ethical, political, technical and economic challenges (Safran et al., 2007). These have been themed by the Agency for Healthcare Research and Quality (AHRQ) into six discrete topics, as detailed in Table 2.5. These topics were utilised as a framework in the design of this research.

Table 2.5: AHRQ data collection challenges **-** The six challenges of today's performance measurement data collection and reporting environment (AHRQ, 2006).

Challenge	**Challenge components**
Inefficiency	• Variations in data collection. • Collection and reporting utilise different taxonomies and data definitions leading to requirements for data validation and continuous updating • Documentation and data quality. • Incomplete clinical documentation. • Disparate electronic systems. • Manual data abstraction. • Inconsistent policies and practice. • Provider staff resources: Increased staffing resources in conjunction with reporting requirements
Variations in measurement systems	• Mandatory vs. voluntary reporting. • Differing reporting formats for different institutions, sometimes for the same disease and patient cohort.
Organisational and cultural issues	• Health care organisations must have stakeholder acceptance, internal change organisation and a culture that allows the continuing provision of reliable data and implementation of changing requirements.
Technological barriers for electronic health records (EHRs)	• Uncoordinated implementation of health IT systems locally and nationally. • Interoperability issues. • Cost of technology. • Lack of understanding of the improvement role that EHRs can play in improving data reporting nationally. • Lack of minimum common data sets for population health and quality measurement. • Security and privacy issues. • Data ownership issues.
Economic pressures	• Costs of collecting data. • Cost of dissemination and interpretation of performance data within organisations.
Competing priorities	• Variations in measure sets, data metrics and taxonomies. • Lack of alignment between the institutions mandating the reporting. • Absence of a national health care quality data set and report card. • Privacy of individuals versus reporting requirements. • Keeping up to date with the changing reporting requirements.

2.8.1 Challenge 1: Inefficiency

The acquisition of data presents a strong challenge for all organisations, particularly when there are variations between the data collected and the minimum data sets required. Data is costly to collect (Cimino, 2011). Inefficiency can have a negative impact on consistent data collection, quality of documentation and data, and on organisations' staffing and financial resources (AHRQ, 2006, HIQA, 2013c), as detailed in Sections 2.8.1.1, 2.8.1.2 and 2.8.1.3 below.

2.8.1.1 Variations in data collection

Variations in taxonomies and data definitions used in collection and reporting requirements can affect the quality of data as well as adding further costs, due to requirements to validate transmitted data and continually update forms and systems as collection metrics change (AHRQ, 2006). Changes in data fields can involve retraining staff and modifying work practices and electronic systems to ensure that proper and consistent data fields are in place across the multiple departments that serve as patient entry points (IOM, 2009). The absence of broadly accepted criteria and standards can result in duplication of effort, increased expense, and lost opportunities to reuse data (AHRQ, 2006, Comber, 2006a, HIQA, 2013c).

Table 2.6: Problems caused by disparate data element definitions and taxonomies (Adapted from (AHRQ, 2006)

Lack of data element definitions and taxonomies by data-seeking organisations causes:

- Time-consuming and complex operations for data acquisition from electronic systems
- Implementation of multiple and disparate systems for data collection within organisations
- Unnecessary barriers to the exchange of health information
- Resource-intensive data-mapping to link systems and reporting requirements
- Conflicts or differences between data sets

There can be variations in data element descriptive fields and submission variables, for example gender (AHRQ, 2006). Some of these variations will be resolved through the introduction of standardised data sets, such as demographics, and also by increased accessibility to minimum data sets and dictionaries currently in use (HIQA, 2013b, HIQA, 2013c).

Efforts are currently underway in Ireland to introduce standard definitions through the publication of data dictionaries in the areas of HIPE and HSE KPIs (HSE, 2011, HIQA, 2012e). In addition HIQA have proposed a national standard demographic dataset for use in health (HIQA, 2013b), as well as principles for data collection, which emphasise the requirement for transparency and collaboration with regard to data collection where the availability of data dictionaries is recommended (HIQA, 2013c).

These variations also impact on the reuse of electronic data and are further discussed in Section 2.8.4.3.

2.8.1.2 Documentation and data quality

Quality data can be defined as reliable data, available in a timely manner to decision-makers (HIQA, 2012e). As secondary health data is being increasingly used, it is important to ensure that the original source data is accurate so as to correctly inform the planning and management of the health services, as discussed in Section 2.4 (HIQA, 2012e). Ideally this data collection would happen at the point of original data capture to be then reused in line with the "collect once, use many times" paradigm, as discussed in Section 2.2, which can be supported by EHRs, as discussed in Section 2.6.1.

The production of quality data for secondary use is impeded by failure to understand the underlying requirements, through errors due to manual data abstraction, and in environments collecting data without consistent policies and practices being in place (Safran et al., 2007). The location of data varies within institutions, as discussed in Section 2.5, and data collectors are faced with challenges including incomplete clinical documentation and extraction of the required information from these disparate locations, which can result in incomplete, biased and unusable data (McDevitt and Comber, 2009). Data collectors require training in areas such as data definitions and their interpretations, as well as protocol registry and data collection systems, to ensure that data are abstracted and reported accurately(AHRQ, 2010). Ensuring data quality and complete clinical documentation requires the cooperation of disparate organisations in order to identify and harmonise the key data required, to ensure integration into clinical workflow (Cusack et al., 2012, Osheroff et al., 2005, AHRQ, 2010).

2.8.1.3 Provider staff resources

The collection of information should not impose a burden on the health system, rather it should be collected as a routine part of the provision of care (HIQA, 2012e). Policy makers must understand the strengths and weaknesses of data already collected and how these data align with the purpose of collection before extending data collections, as the collection of these data presents challenges (HIQA, 2013c, GAO, 2010) and is costly (Cimino, 2011). HIQA and the AMIA recommend investigation of the feasibility of data collection and an evaluation of the potential burden on the data provider to collect the data prior to implementing new data requirements for data reporting (Cusack et al., 2012, HIQA, 2013c, AHRQ, 2010).

Increased levels of staffing often directly result from new, or changed, reporting requirements due to the increased demand and also the variations as discussed in Section 2.8.1 (Wilkinson et al., 2007, Pham et al., 2006, Bradley et al., 2006). Inadequacy of IT systems is an underlying cause of the staffing burden associated with reporting (Pham et al., 2006). EHRs allow enhanced data collection and reduce the reporting burden, as discussed in Section 2.6.1, (Cusack et al., 2012), but it is still not uncommon to have departments dedicated to data management, including record abstraction, resulting in duplication and redundancy in the process of the actual collection, despite the existence of electronic data (Cimino, 2007). Work process review must be considered to improve efficiency (Kellermann and Jones, 2013).

2.8.2 Challenge 2: Variations in measurement systems

Data collection in Ireland, similar to other countries, has evolved over time. Lack of coordination has resulted in variations in quality as well as duplication (HIQA, 2013c). The data required by different management systems can vary, even when reporting on the same disease; for example the symptomatic breast KPIs as required by the NCCP and HIQA (HIQA, 2007, NCCP, 2010).

Variations induced by the mandatory or voluntary status of data reporting is not significant in Ireland, as there is little mandatory health data reporting, with some exceptions such as reporting of specified infectious diseases (Reilly, 2011, HIQA, 2013a).

2.8.3 Challenge 3: Organisational and cultural issues

Delivery of government policies, such as the requirement for data reporting, can be jeopardised when policy decisions are separated from delivery responsibilities (James, 2003). Implementation of these policies may require transformation of the culture of public services through the application of project management disciplines, managed by project owners with an understanding of the coalface implementation setting (Wanna, 2007). One way forward is the appointment of “senior responsible owners”, who are responsible for monitoring the policy implementation at local level and anticipating and resolving challenges to that implementation (Wanna, 2007).

This is also applicable in the context of implementation of ICT programmes where leadership, both national and local, has been shown to be one of the fundamental requirements for successful implementation of EHRs and other clinical systems (Kellermann and Jones, 2013, Wilkinson et al., 2007, Osheroff et al., 2005, Cresswell et al., 2011, Callen et al., 2008).

Staff training is essential to ensure that EHRs are used and managed effectively, and to provide continuity (Gelmon and Droppers, 2008, Kellermann and Jones, 2013, AHRQ, 2010). Education of staff with regard to the objectives of data collection is also important to engender the reporting process to clinicians (McDevitt and Comber, 2009, Wilkinson et al., 2007).

The factors that determine how well organisations can adjust to externally imposed requirements for the purpose of providing reliable data include stakeholder acceptance, flexibility of internal structures and the organisational culture (Safran et al., 2007, Kellermann and Jones, 2013). Institutions providing data require robust change management structures and communication pathways, both internal and external, to meet the changing requirements placed upon them by accrediting and standards bodies (Safran et al., 2007). Two-way communication is essential to ensure that policy challenges are recognised early in the process and acted upon, in addition to ensuring that the policy implementation remains aligned to the national policy requirements and meets its objectives and full potential (Wanna, 2007).

2.8.4 Challenge 4: Technological barriers for EHRs

2.8.4.1 Identifiers

Healthcare identifiers are key building blocks essential for the progression of electronic health (eHealth) (NEHTA, 2012, HIQA, 2009, HIQA, 2011c). They allow the identification and association of all relevant health information about an individual. They are essential in enabling the effective implementation of an EHR and in supporting secondary use of health data (DOH, 2004). The identifiers for individuals, organisations and healthcare practitioners, as proposed for Ireland (DOH, 2004, HIQA, 2009, HIQA, 2011c) are not dissimilar from other countries, as is evidenced by the similarities with those of Australia, as detailed in Table 2.7.

Table 2.7: Comparison of health identifiers in Australia and Ireland (HIQA, 2009, HIQA, 2011c, NEHTA, 2012). The Irish health identifiers are proposed.

Identifier group	**NEHTA**	**HIQA proposal**
Individuals	Individual Healthcare Identifier (IHI) – Allocated to all individuals enrolled in the Medicare program or those who are issued with a Department of Veterans' Affairs (DVA) treatment card, and others who seek healthcare in Australia.	IHI – recommended to be allocated to all people in Ireland enabling health and social care to be delivered to the right patient, in the right place and at the right time.
Healthcare practitioners	Healthcare Provider Identifier – Individual (HPI-I) - Allocated to healthcare providers involved in providing patient care.	Healthcare Practitioner Identifier (HPI) – recommended to be allocated to all healthcare practitioners. Its purpose is to identify the individual as one and the same person and to allow the "attaching" of other information (such as name, address, and contact details) to them.
Healthcare organisations	Healthcare Provider Identifier – Organisation (HPI-O) - Allocated to organisations (such as a hospital or medical clinic) where healthcare is provided	Healthcare Organisation Identifier (HOI) - A healthcare organisation identifier is a unique, non-transferable number assigned to healthcare organisations in Ireland. It will allow the attaching of a dataset to identify its location, contact details and operational sites.

The Irish Data Protection Commissioner (DPC) also supports the establishment of an individual health identifier, which is not the Personal Public Service Number (PPSN), as the use of the PPSN is deemed inappropriate and would expose the public to invasions of their privacy (Davis, 2009). The introduction of the Health Information Bill has been on the DoH agenda since close of consultation in 2008 (DOH, 2013b).

There are data items in the selected case, additional to those described above, which also require the use of standard identifiers. They are the drug protocols and the drugs themselves. It is well recognised that there is a requirement for a standard for representing medicines in information systems and electronic communications (NHS, 2013, Kernan, 2012, HIQA, 2012a). Ireland does not yet have a national drugs dictionary. The PCRS utilises a drug database to support their reimbursement services, but it does not generally include hospital-specific items and is not publically available (HIQA, 2013a).

2.8.4.2 Data standards and interoperability

The International Standards Organisation (ISO) define levels of interoperability as functional or semantic, where semantic interoperability allows systems to exchange information without a requirement for interpretation (ISO, 2005). Standards have been defined as “a definition or format that has been approved by a recognised standards organisation or is accepted as a de facto standard by the industry” (HiMSS, 2010), and are required for clinical vocabularies, message exchange and ontologies (McLean, 2006).

These two components, data standards and interoperability, comprise a framework that allows communication between disparate ICT systems, which is fundamental to achieving data aggregation (Karp et al., 2008, Stead et al., 2005, empirica GmbH, 2008, Marcheschi et al., 2005) as well as being key to eliminating the information silos generated by legacy and non-integrated health care systems (Barton et al., 2011, Cimino, 2011, McDonald, 1997).

It must be recognised that many information systems currently in use have been in place prior to the implementation of standards for information exchange and therefore may not be capable of exchanging data electronically within institutions, much less with external organisations (Barton et al., 2011). In the absence of interoperability, electronic data extraction and reporting can be facilitated by manual mapping of data (Ancker et al., 2011), however, data mapping requires resources (Rea et al., 2012, El Fadly et al., 2011, Hernandez et al., 2009).

The requirement for eHealth interoperability standards is well recognised in Ireland. The HIQA consultation document on Irish national health interoperability standards recognises the deficiencies of current systems, in addition to identifying the requirement for key building blocks, based on international standards, to be in place to build a robust eHealth infrastructure (HIQA, 2011a). Managing organisations of national health data collections are now encouraged to confer with HIQA to determine suitable standards in the design of new data collections or upgrades of current systems (HIQA, 2013c). Economic pressures have thus far prevented the adoption of the Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) as an Irish national standard (HIQA, 2011a).

2.8.4.3 Lack of minimum common data sets

A minimum data set has been defined as “the minimum set of data elements that are required to be collected for a specific purpose” (HIQA, 2012e). Nationally agreed and approved minimum data sets ensure that the same data can be reused for multiple purposes without a requirement for new data collection or interpretation (Barton et al., 2011).

A recent development in Ireland has been the production of a proposed national demographic dataset for health which will help to consistently identify demographics (HIQA, 2013b). In addition to this there is also a requirement to define and harmonise clinical data content standards which can support both patient care and secondary data uses in the areas of disease surveillance and population health (Barton et al., 2011). Currently, there are a limited number of national minimum datasets produced for specific disease indications in Ireland (HIQA, 2007).

The requirements of public health must be considered when designing minimum data sets so as to maximise the usability of the data collected (Barton et al., 2011). Cooperation between agencies collecting data is recommended to reduce and remove duplications and develop greater efficiencies (HIQA, 2013c).

2.8.5 Challenge 5: Economic pressures in Ireland

Economic pressures on hospitals and staff moratoriums lead to competing priorities for both financial and personnel resources (Burke, 2010). The increasing requirement for additional personnel to facilitate data reporting, as detailed in Section 2.8.1.3, will add to these pressures. Ensuring that data is available to be collected and reusing data in so far as possible will minimise this resource requirement (HIQA, 2013c). No personnel resources have been provided to facilitate the data collection in the selected case (O'Reilly, 2013).

Since 2009, publically funded Irish hospitals no longer have the authority to procure ICT systems without prior financial approval from the Centre for Management and Organisational Development (CMOD). This department is responsible for monitoring and approving ICT expenditure in civil and public service bodies with an emphasis on national solutions (HSE, 2012a, Duffy, 2009).

2.8.6 Challenge 6: Competing priorities

This section will focus on the competing priority of the privacy of an individual with the requirements of the health services. A number of the other competing priorities are identified by the AHRQ, as detailed in Table 2.5, such as staffing resources and variations in measurement systems which have been included in the relevant sections above.

2.8.6.1 Privacy of individuals / Data stewardship

Legislation on personal health information (PHI) is essential to underpin the governance and safeguards that are required to ensure that PHI is stored, transmitted, accessed and used correctly (Longstaff, 2005, Bloomrosen and Detmer, 2008). The introduction of such legislation and a framework for data stewardship, as discussed in section 2.4 and detailed in Table 2.8 below, would balance the rights of an individual to have their PHI protected, and their wish to have an improved health service (Bloomrosen and Detmer, 2008).

Table 2.8: Data stewardship principles (Bloomrosen and Detmer, 2008)

Data stewardship principles:
- Accountability, including governance, oversight, and the application of relevant regulations to the appropriate extent and level.
- Transparency, including clearly understood policies and procedures regarding data structure, processing, and delivery of data, and business processes and practices.
- Notice to patients and other legitimate users.
- Technical issues, e.g., data security, and quality, de-identification, and costs of re-identification.
- Patient consent of appropriate granularity.
- Permitted uses and disclosures including data aggregation and analyses.
- Enforcement and remedies.

Ireland is awaiting the publication of its first Health Information Bill, on which public consultation closed in September 2008 (DOH, 2013b). In the absence of specific legislation, Irish health data governance falls under the remit of the Data Protection Commissioner (DPC) and Data Protection Act 1988 and its amendments (Data Protection Commissioner, 2013b). Each healthcare provider has an appointed “data controller” who is required to be registered with the DPC, and is the person legally responsible for the control and content of the patient data. This responsibility includes ensuring that data is fairly obtained and is shared only for the purposes for which consent has been received, in addition to ensuring that individual patient consent is given in circumstances where data is to be shared beyond that of its original use (Data Protection Commissioner, 2013a). There are four main HSE data controllers, in addition to individual data controllers for each voluntary and private hospital (Data Protection Commissioner, 2013c).

Aside from legislation there is a requirement to gain the public’s trust and support for the secondary use of data through education and awareness of the societal benefits of secondary use (Bloomrosen and Detmer, 2008, Longstaff, 2005).

2.9 Identification of Stakeholders

The identification of stakeholders is multifaceted and essential to ensure that all aspects of the research, or project, are articulated. Stakeholders have been defined as "any group or individual who can affect or is affected by the achievement of the organization's objectives" (Freeman, 1984). The Prince2^{TM} project methodology identifies three basic types of stakeholder, as indicated in Table 2.9, whose interests are crucial to the success of a project (OGC, 2009).

Table 2.9: The three project interests as defined by Prince2TM (OGC, 2009)

Stakeholder	Viewpoint
Business user	Individuals or organisations who require the project to meet a business need
User	Individuals or groups who will use, support or be impacted by the project outputs
Supplier	Individuals or groups who will produce the project outputs

Identification of stakeholders involves consideration of the three project interests. Known stakeholders can help identify other potential stakeholders. Stakeholders are also critically placed to assess the implementation of the project and identify challenges to its implementation (PRINCE2 learning provider ILX Group, 2013).

2.10 Key Findings of the Literature Review

As outlined in Section 2.4, Ireland, similar to other countries, has increasing requirements for health data, driven by the requirements of a variety of institutions. Data is required to underpin safe and effective health services, and to allow surveillance and audit by public health bodies to ensure that the health services are effective and accessible to all within the allocated budget. This requirement includes the domain of cancer control where the increasing prevalence, complexity and cost of disease treatment are generating a growing need for timely, accurate and applicable data.

Secondary use of health data is a complex phenomenon which faces challenges including variation in measurement systems, organisational, cultural, technical and economic issues, in addition to competing priorities, as discussed in Section 2.8. The burden created by data collection requirements, particularly where these data must be manually extracted from disparate sources, is well recognised. Data collection programmes require implementation in a manner that is cognisant of economic pressures and staffing resources and in so far as possible should avoid creating additional burdens for the institutions involved. The programmes must also consider the requirement to ensure that the data required are obtained fairly and in accordance with relevant legislation and best practices with regard to privacy and the use of personal health information.

ICT and EHRs have been shown to have the potential to increase the efficiency of both data collection and reporting through the reuse of electronic health data, as discussed in Section 2.6.1, thus potentially reducing personnel resource requirements. This reuse of electronic data is dependent on the implementation of legislation, data standards and a robust ICT infrastructure. Initiatives are underway in Ireland to construct the building blocks to improve the secondary use of health data through the implementation of standards, the introduction of the required legislation, and promotion of cooperation amongst requesters of data.

Improvement in technology is not the only factor that will facilitate an improvement in data collection. The sociocultural challenges associated with change management and programme implementation are equally important. A well-conceived and structured project management framework, with responsible local and national owners, and bilateral communication can secure stakeholder buy-in and achieve broad acceptance of organisational aims and means, as discussed in Section 2.8.3.

2.11 Conclusion

This chapter explored the relevant and scholarly literature focused on the requirement for health data for secondary use, the challenges, and facilitators to its collection, as well as on the setting of the selected case. The challenges to the secondary use of health data have been themed, by the AHRQ, into six discrete topics. These themes were utilised by the researcher as a framework in the design of this research, as described in Chapter 4.

Chapter 3. The Selected Case

3.1 Background

Ireland has multiple providers of health care, both in the private and public service. The HSE is responsible for the provision of publically funded Irish healthcare, including those services in voluntary hospitals (HSE, 2012d). Irish healthcare is undergoing a number of organisational changes that have precipitated the implementation of the selected case. These include the introduction of universal health insurance, which is underpinned by the ‘money follows the patient’ approach to funding (HSE, 2013b).

Historically, drugs administered in secondary care, including systematic anti-cancer therapy (SACT), were funded as an integral part of each hospital’s overall budget. SACT will require a different funding stream from other drugs in the “money follows the patient” funding model due to their high cost, as discussed in Section 2.3.1 (DOH, 2013a).

The National Cancer Control Programme (NCCP), in anticipation of the required change to the funding model, proposed that the funding stream for secondary care SACT be disconnected from the overall secondary care budget and funded through direct reimbursement to the hospitals based on individual patients’ consumption of these medications (HSE, 2012e). This would facilitate the “money follows the patient” model in addition to centralising the SACT national budget, as discussed in Section 2.3, and would ultimately allow the NCCP to implement many of the core components of a cancer control programme, such as clinical audit (O'Reilly, 2013). The data required to support such a scheme is available in other jurisdictions in national cancer information systems, as discussed in Section 2.4.1.

3.2 The Selected Case

In the absence of a national cancer information systems (NCIS), to provide the required data, the NCCP established the SACT reimbursement system as a new national data collection. This data collection and reimbursement is facilitated through the Primary Care Reimbursement Service (PCRS) and requires collation and manual entry of a minimum dataset via a specifically designed web portal (NCCP, 2012b). The patient-specific data reported by the hospitals include the patient’s diagnosis, the treatment protocol used and the drugs administered. This reporting is intended to guide the funding allocated to individual hospitals for the treatment of patients with cancer, in addition to facilitating the collection of treatment numbers and monitoring adherence to national protocols (O'Reilly, 2013).

The SACT reimbursement system was implemented in 2012 for a limited number of drugs, with the intention of including all newly approved secondary care SACTs, in addition to extending the data collection to other high-cost SACTS, and eventually all SACTS (O'Reilly, 2013). While no personnel resources were allocated to the hospitals to facilitate the collection and reporting of these data, financial support was provided to the hospitals to register existing patients (O'Reilly, 2013).

3.2.1 Data input processes of the SACT reimbursement scheme

The steps in the process of data entry to the web portal are detailed in Table 3.1. Prior to data entry, the data must be located, verified and collated. Following submission of the claim, it must be tracked to confirm payment and matched to incoming financial transactions. The aspects of data collation and invoice reconciliation of the selected case are beyond the scope of this research.

Table 3.1: SACT reimbursement scheme data entry steps

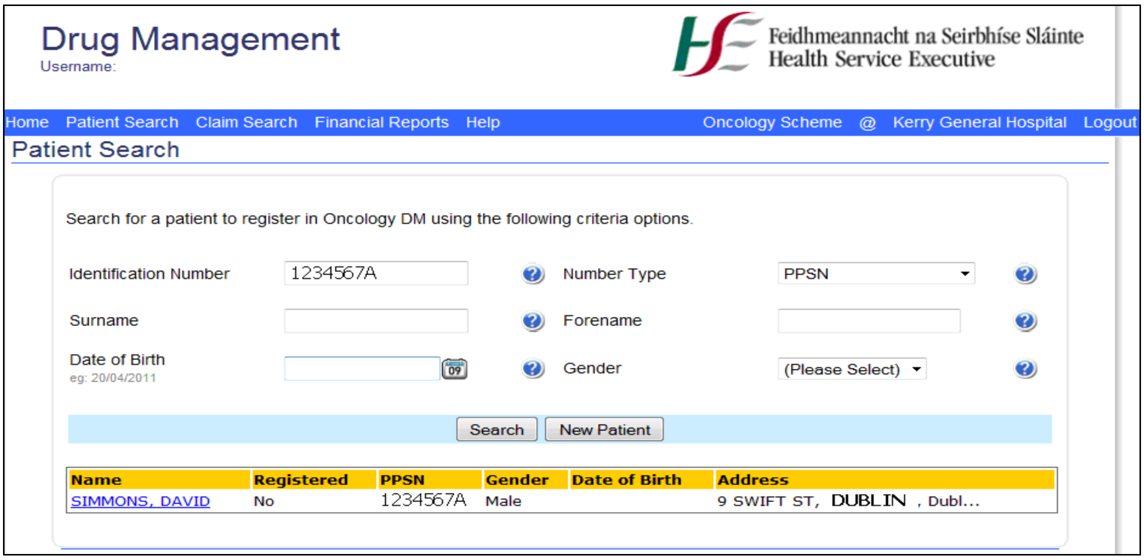
Step	New patient	Previously registered patient
1	Login to web application	
2	Register patient to the SACT reimbursement scheme	Search for patient on web application
3	Make a claim	

3.2.2 Data input screens of the SACT reimbursement scheme

Access to the SACT reimbursement scheme is via the internet. There are a number of different screens utilised when registering patients and making reimbursement claims, as detailed in the following sections.

3.2.2.1 Patient search and patient registration

The SACT reimbursement scheme data collection is a subset of the wider PCRS database and allows reuse of existing PCRS patient demographic information in the registration of patients so as to minimise data entry. Patients not previously registered must be registered as new patients, and the patient data entered manually. The patient search can be conducted using a combination of fields, or one of the patient identification numbers. Figure 3.1 is a sample of this screen with test patient details displayed.



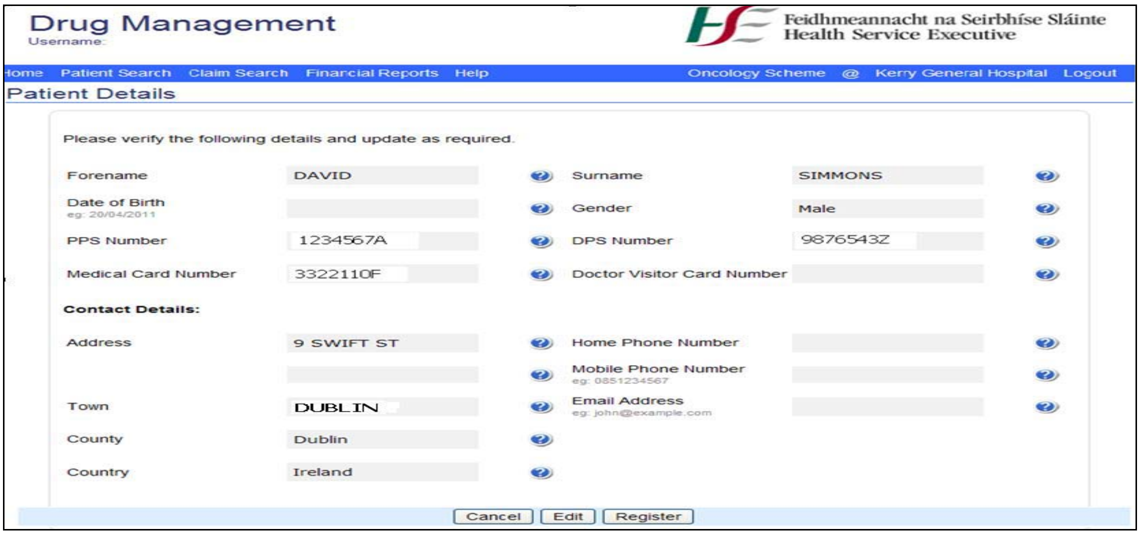
Search for a patient to register in Oncology DM using the following criteria options.

Identification Number: 1234567A
Number Type: PPSN
Surname:
Forename:
Date of Birth (eg: 20/04/2011):
Gender: (Please Select)

Name	Registered	PPSN	Gender	Date of Birth	Address
SIMMONS, DAVID	No	1234567A	Male		9 SWIFT ST, DUBLIN , Dubl...

Figure 3.1: Web portal patient search screen

Once a patient is selected, their information is displayed for verification, as shown in Figure 3.2. The patient detail can now be registered to the SACT reimbursement scheme, or the user may edit the information or cancel the process.



Please verify the following details and update as required.

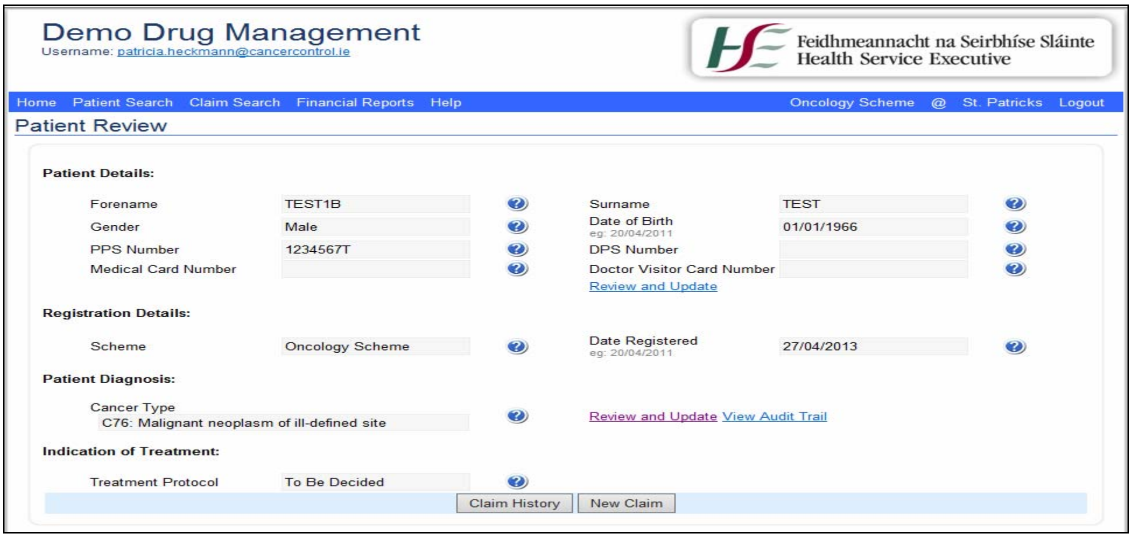
Forename: DAVID
Surname: SIMMONS
Date of Birth (eg: 20/04/2011):
Gender: Male
PPS Number: 1234567A
DPS Number: 98765432Z
Medical Card Number: 3322110F
Doctor Visitor Card Number:

Contact Details:

Address: 9 SWIFT ST
Home Phone Number:
Mobile Phone Number (eg: 0851234567):
Town: DUBLIN
Email Address (eg: john@example.com):
County: Dublin
Country: Ireland

Figure 3.2: Web portal patient details screen

Other screens allow the user to review the patient information and update the data where necessary. An audit trail is maintained of these changes; updates to the cancer type require entry of the reason for change and may be viewed as an audit trail log. The application does not currently support the registration of patients with multiple cancers.



Demo Drug Management
Username: patricia.heckmann@cancercontrol.ie

Feidhmeannacht na Seirbhíse Sláinte
Health Service Executive

Patient Review

Patient Details:

Forename: TEST1B
Surname: TEST
Gender: Male
Date of Birth (eg: 20/04/2011): 01/01/1966
PPS Number: 1234567T
DPS Number:
Medical Card Number:
Doctor Visitor Card Number:

Registration Details:

Scheme: Oncology Scheme
Date Registered (eg: 20/04/2011): 27/04/2013

Patient Diagnosis:

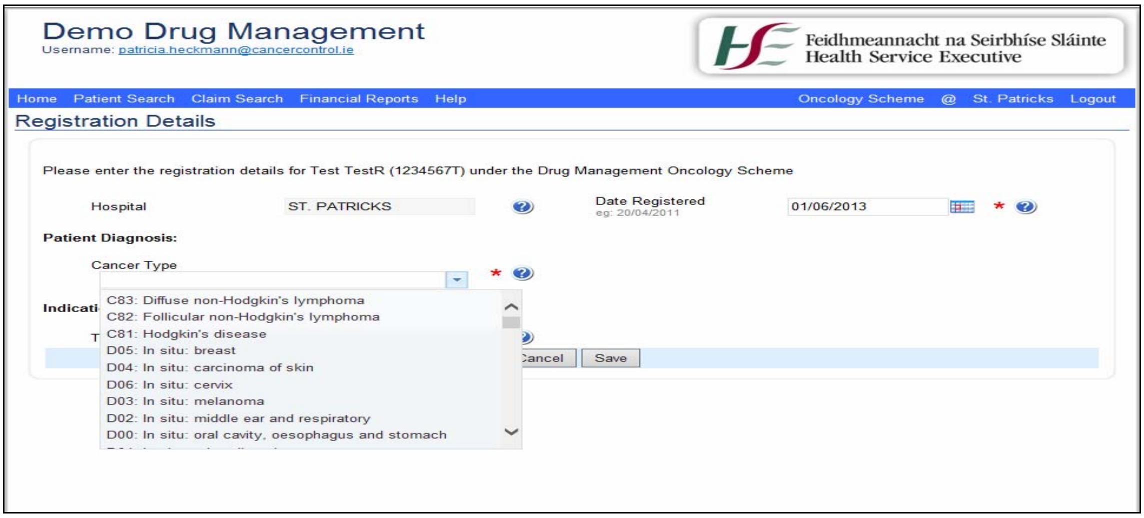
Cancer Type: C76: Malignant neoplasm of ill-defined site

Indication of Treatment:

Treatment Protocol: To Be Decided

Figure 3.3: Web portal patient review screen

Figure 3.4 displays the patient registration screen where the patient cancer type and treatment protocol can be selected from drop-down lists. Saving these data completes the registration of the patient to the SACT reimbursement scheme.



Demo Drug Management
Username: patricia.heckmann@cancercontrol.ie

Feidhmeannacht na Seirbhíse Sláinte
Health Service Executive

Registration Details

Please enter the registration details for Test TestR (1234567T) under the Drug Management Oncology Scheme

Hospital: ST. PATRICKS
Date Registered (eg: 20/04/2011): 01/06/2013 *

Patient Diagnosis:

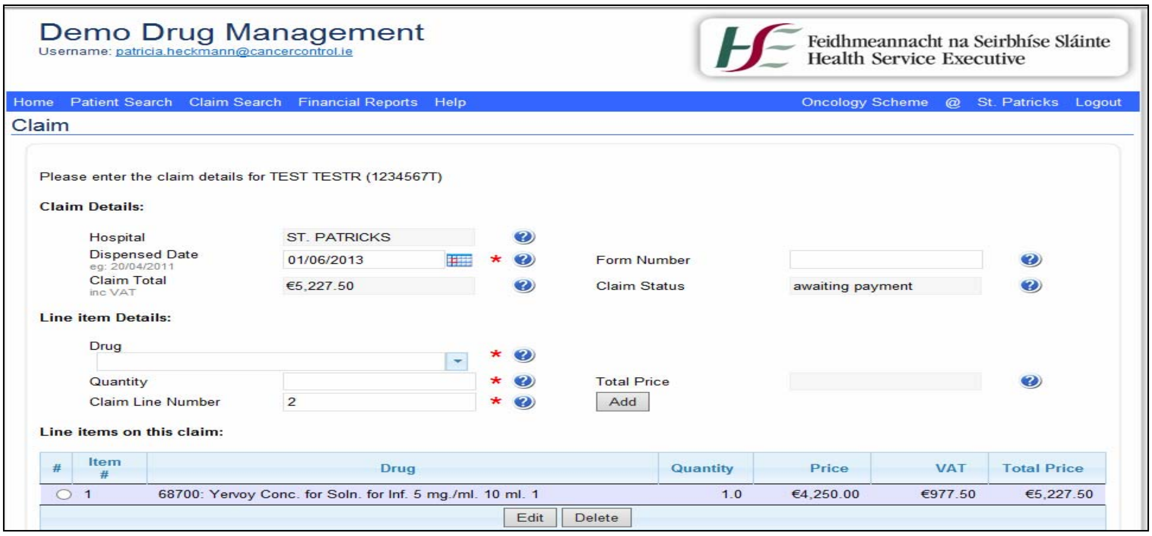
Cancer Type *

C83: Diffuse non-Hodgkin's lymphoma
C82: Follicular non-Hodgkin's lymphoma
C81: Hodgkin's disease
D05: In situ: breast
D04: In situ: carcinoma of skin
D06: In situ: cervix
D03: In situ: melanoma
D02: In situ: middle ear and respiratory
D00: In situ: oral cavity, oesophagus and stomach

Figure 3.4: Web portal clinical details registration screen

3.2.2.2 Submitting a reimbursement claim

Claims for reimbursement may only be made for registered patients. This is done in the claim screen as displayed in Figure 3.5. Details of the number of units of drug that are administered to a patient are entered. Multiple lines may be added to the claim. There are also options to amend, delete or cancel the claim in similar screens, in addition to reports on claims, and a payment log.



Username: patricia.heckmann@cancercontrol.ie

Please enter the claim details for TEST TESTR (1234567T)

Claim Details:

Hospital: ST. PATRICKS
Dispensed Date (eg: 20/04/2011): 01/06/2013 *
Claim Total (inc VAT): €5,227.50
Form Number:
Claim Status: awaiting payment

Line item Details:

Drug: *
Quantity: *
Claim Line Number: 2 *
Total Price:

Line items on this claim:

	#	Item #	Drug	Quantity	Price	VAT	Total Price
○	1		68700: Yervoy Conc. for Soln. for Inf. 5 mg./ml. 10 ml. 1	1.0	€4,250.00	€977.50	€5,227.50

Figure 3.5: Web portal claim screen

3.3 Conclusion

This chapter describes some of the reasons for the introduction of the SACT reimbursement scheme. The principal ones being centralisation of the SACT national budget and facilitation of the government's "money follows the patient" policy. It also familiarises the reader with the data input processes of the SACT reimbursement scheme and provides a background for the research steps taken to answer the second research question, relating to the personnel resources required to input the data in the selected case, which are detailed in Chapters 4 and 5.

Chapter 4. Research Methodology

4.1 Introduction

This chapter outlines the methodology employed by the researcher while investigating the research questions, presented in Chapter 1, as to the challenges and facilitators to the collection of health data for secondary use in Ireland and the personnel resources required to input the data of the selected case. It presents the rationale underpinning the choice of methodology, the data collection tools used, the framework for data analysis, as well as ethical and validity considerations in addition to reliability issues.

The researcher's intention, at all times, was to review and collate all materials gathered in a holistic way and not as a critical analysis of the selected case. The fundamental purpose was to discover the challenges and facilitators, as they exist, and to further the understanding of the phenomenon through conclusions and recommendations.

4.2 Role of Literature Review in Research

The role of the literature in answering the research questions and meeting the research aims are described in Chapter 2. The main theme explored was that of the secondary use of health data in the domain of public health and cancer control. The review also provided the researcher with a framework to select the case to be studied and to present the research findings.

4.3 Choice of Research Design and Strategy

Research design encompasses the research framework from proposal to analysis and can be viewed as the juncture of philosophy, strategy of inquiry and research methods (Creswell, 2009). The intention of this research was to explore and describe the challenges and facilitators to the complex phenomenon of health data collection for secondary use. The philosophy of this research is one of constructivism and pragmatism, as it is seeking to understand and demonstrate these challenges and facilitators, as described by multiple stakeholders, rather than to test a theory, and also to examine the practical real-world topic of resource implications in the setting of the selected case (Creswell, 2009).

This type of research requires assimilation of narrative and quantitative data from the data sources of the research, such as documents, interviews and questionnaires. It is neither experimental nor survey research, both of which provide outcomes that can be statistically analysed. The research questions are "what" questions, of an exploratory nature, seeking to develop an in-depth description of the phenomenon (Yin, 2009). Having evaluated the various approaches of qualitative research the researcher concluded that case study strategy best matched the proposed research. The selected case allowed an opportunity to further analyse one subunit of the challenges and sought to explore and quantify the personnel resources necessary to input the required data. The researcher considered that this would be best facilitated by a quantitative survey research strategy, as the subunit analysis is not experimental (Creswell, 2009).

The resultant selected strategy for the enquiry was an embedded single-case design case study, where the embedded subunit research further explored one aspect of the primary research, and the qualitative and quantitative data collected simultaneously.

4.3.1.1 The role of the researcher and ethics

Qualitative research involves the researcher in a way that may introduce ethical, personal and strategic issues into the process (Locke et al., 2007). To this end, researchers may explicitly identify their potential biases, values and personal backgrounds, which might influence their choice of research in addition to interpretations formed during the research.

The researcher’s pragmatic constructive worldview, which perceived the case as a missed opportunity for the reuse of existing electronic health data for secondary purposes, may have influenced the direction of this research (Creswell, 2009). The researcher, being aware of this potential limitation, took steps to ensure that the data collection methods were reliable and valid, so as to ensure control of these biases through careful design and data validation (Richards, 2008).

Ethics approval was sought prior to commencement of the research, as cooperation and participation of hospital staff, system vendor staff and HSE employees was required. This was received from the School of Computer Science & Statistics, University of Dublin, Trinity College, in December 2012, and is included as Appendix 1. “Gate keeper” permission, included as Appendix 2, was sought from the director of the NCCP, to allow access to the setting and to secure permission to engage with the NCCP participants, in addition to ensuring that there was full knowledge of the research being undertaken. All invitations to participate were accompanied by an information leaflet specific to that phase, which summarised the research proposal. The participant consent forms and information leaflets are detailed in Appendix 3.

4.3.1.2 Case selection and study design

Case studies may utilise a single-case or multiple-case study design, within which there may be subunits (Yin, 2009). The study of small numbers of cases, sometimes just a single case, allows for the collection of data on a large number of features of each case (Gomm et al., 2000), with embedded case studies allowing the researcher to devote attention to one or more subunits within the case. The rationale selected for the purpose of this research was that of a "representative" or "typical" case. The aim was to capture the contemporary challenges, and facilitators, to the phenomenon of health data collection for secondary use, as identified in the course of the literature review in Chapter 2, and to be further illuminated by the experience of the stakeholder interviewees, as described in Chapter 6. The outcomes of this research can then be assumed to be informative about the challenges to the secondary use of data as experienced by the average person or institution (Yin, 2009). Therefore, the selection of the case is crucial: it must have strategic importance in relation to the general problem, and be selected on the basis of the information that it can be expected to provide.

The researcher wished to examine a "typical" case which could reveal the challenges and facilitators to the collection of health data for secondary use. Ideally, the case would be one that was not already an embedded practice, which had been recently introduced and was to be extended. The rationale for these requirements was to gather data from stakeholders engaged in the reporting of these data at a time when the challenges were still very apparent and the pattern of reporting was yet to be fully established.

The researcher, a pharmacist with an understanding of both hospital and NCCP requirements, had knowledge of a case that suggested itself as "typical" and fell within the domain of the research. The researcher then analysed the subunits of the case, represented as potential challenges, aligned to the challenges elicited from the literature review to ensure validity in the case selection. The details of the case and the subunits are detailed in Table 4.1.

The case suggested by the researcher's personal experience was the SACT reimbursement scheme, as described in Chapter 3, which had been introduced in 2012. This scheme required data to be reported, by the institutions collecting the data in the course of the clinical care of the patient, for the purpose of reimbursement. The reporting of these data is not mandatory, but dependent on voluntary reporting by the organisations. The intention is to extend the scheme to include additional SACTS, thus increasing the amount of data to be reported. The selected case also presented an opportunity to quantify the personnel resources necessary to input the required data. This was one of the reported challenges from the literature review, in addition to being perceived as a significant challenge by the researcher.

Table 4.1: Selection of the case. Subunit analysis was conducted of the selected case to determine its suitability as the selected research case.

Subunit/Challenge	**Potential to exist in selected case**	**Potential challenge**
Inefficiency	Yes	Incomplete data sets, duplicate data entry, manual data entry, no quality data checks, dispersed location of complete data set between electronic systems and paper charts.
Variations in measurement systems	Yes	New data requirement: different yet similar to other data being collected for NCRI, NCCP.
Organisational and cultural issues	Yes	New reporting requirement, potential work practice change.
Technological challenges (EHRs)	Yes	Disparate/legacy systems, lack of legislation, use of standards, identifiers, system interoperability and data sets. Issues with regard to “data ownership”. Security and privacy issues.
Economic pressures	Yes	No additional resources for staffing, ICT.
Competing priorities	Yes	Data protection: Patient-specific information to be communicated outside institution collecting as a primary data source. Local vs. National priorities. Clinical vs. Administrative priorities.

4.3.2 Benefits of case study research

Case studies, contrary to experimental studies, are generally the preferred method when asking an exploratory “what” question; the researcher has little control over events and the focus is on a contemporary complex phenomenon (Yin, 2009). The phenomenon being explored was the secondary use of health data; the case study structure allowed the researcher to take a holistic approach and rigorously explore the phenomenon in context, using a variety of data sources, such as literature, interviews, direct observation and questionnaires, in order to converge the data to illuminate the case and to answer the research questions (Baxter and Jack, 2008, Yin, 2009). The collection and comparison of these data enhances data quality, based on the principles of idea convergence and the confirmation of findings (Baxter and Jack, 2008). The selection of subunits for analysis allows the collection of clear data as well as focussing the case study inquiry. However, the focus of the research must return from the subunit to the wider research in order to avoid the challenges to the secondary use of health data becoming the context rather than the target of the research (Yin, 2009).

4.3.3 Limitations of a case study

The case study was designed to overcome the general case study limitations of lack of rigor and generalisability (Baxter and Jack, 2008, Yin, 2009).

Data was gathered from interviews, questionnaires, a time study and a literature review. The maintenance of the chain of evidence of the gathered data was prioritised and the researcher used a case study database to record the data, as detailed in Sections, 5.2, 5.3 and 5.6. The gathered data from the multiple sources of evidence was analysed and integrated in order to ensure reliability, construct validity, provide data validation and to eliminate bias (Yin, 2009).

4.4 Identification of Stakeholders

The researcher, in line with Prince2 theory as discussed in Section 2.9, identified three primary stakeholder groups: the hospitals, the NCCP and the PCRS, as detailed in Table 4.2 below, in addition to secondary groups such as HIQA, HSE ICT, vendors of relevant software systems, hospital data managers and hospital system integration specialists. Primary stakeholders included those directly involved in the data input, management or implementation of the SACT reimbursement scheme. Secondary stakeholders were defined as those with a potential interest in the data or the method of collection in the broader context of public health and the secondary use of health data. The researcher also anticipated that additional stakeholders might become evident during the course of the semi-structured interviews. The full list of identified potential stakeholders and those interviewed is included in Appendix 16.

Table 4.2: The three project interests (OGC, 2009)

Stakeholder	Viewpoint
Business user - NCCP	Individuals or organisations who required the project to meet a business need
User - PCRS	Individuals or groups who will use, support or be impacted by the project outputs
Supplier - Hospitals	Individuals or groups who will produce the project outputs

4.5 Research Methods

4.5.1 Qualitative - interviews

The interviews were conducted to gather data on various topics that would facilitate the exploration of the challenges to the secondary use of health data in terms of the research aims, as detailed in Table 4.3, and to enable generalisation from the domain of cancer control and the selected case to the wider area of secondary use of health data.

Table 4.3: Research objectives of the semi-structured interviews

Research method	**Interviews of primary and secondary stakeholders (Section 4.5.1)**
Research aim	**Research objectives**
To investigate the selected case in the areas of the challenges and facilitators to the collection of health data for secondary use	To determine the contemporary challenges to the secondary use of health data in the areas of: • Cancer control and the wider domain of public health • The SACT reimbursement scheme • The participants awareness of: - The necessity to collect data for public health and cancer control purposes - Standards used in the domain of health data • To determine potential ways to improve the secondary use of health data • To reveal other potential stakeholders for the purpose of including them in the research • To enable generalisation from the domain of cancer control and the selected case to the wider area of secondary use of health data

The researcher employed a semi-structured interview format to obtain descriptions of the real world of the interviewees in order to interpret the phenomena of the challenges to the secondary use of health data (Kvale and Brinkmann, 2009). Structured interviews were deemed unsuitable for the purpose of this research as questions would have been restricted to a fixed list, some of which might not be relevant to a particular interviewee. The semi-structured format allowed for flexibility in terms of the topics discussed, in accordance with the experience and expected contribution of each interviewee, in addition to allowing the researcher to frame the interviews so as to ensure that the conclusions would be generalisable. Aside from the SACT reimbursement scheme challenges, other topics discussed included the interviewees' understanding of the requirement for national collection of health data generally, as well as specifically relating to cancer. To ensure generalisability, the challenges to the broader phenomena of secondary use of health data and interviewees' views on approaches to improve national secondary use of data were also discussed. The topic list for the semi-structured interviews and the data collection tool are included in Appendix 14 and Appendix 15 respectively.

The researcher’s intention was to conduct the interviews face-to-face and otherwise by telephone, to take notes at all interviews and record interviews when permitted. Notes were confirmed with the interviewees prior to interpretation, to ensure validity and to remove bias. The interview notes were then documented in a case study database to maintain a chain of evidence and to analyse themes, as discussed in Section 5.6. The case study database was also used to record the target interview participants, their domain area, the location of the interview and if it was recorded. Coding of participants was maintained in a separate database to ensure anonymity.

Selection of representative stakeholders, for the purpose of the interviews, relied on opportunistic sampling and availability. A minimum of two stakeholders from each of the three defined primary stakeholder groups were interviewed to ensure representation of each group’s unique challenges. Due to the time constraints and the pragmatic focus of the research representatives of all the identified secondary stakeholder groups identified were not interviewed.

4.5.2 Quantitative – questionnaires

One component of the quantitative phase of the research involved two online self-administered questionnaires, a hospital questionnaire and a software systems vendor questionnaire, both consisting of closed questions intended to answer the research aims, as detailed below in Table 4.4. Self-administered questionnaires have the advantage of being familiar, in addition to being low cost. Disadvantages include lack of truthfulness, in addition to some respondents failing to answer some or all of the questions (Fink, 1993). The 26 centers involved in the administration of SACT were invited to participate in the research to establish a national overview. The three vendors of the disparate SACT preparation systems in use, which were selected for further research as a potential source of required data in an electronic format, were also invited to participate. Screen shots of the questionnaires are included in Appendix 9 and Appendix 10 respectively.

The main aims of the hospital and vendor questionnaires are detailed in Table 4.4. In the case of the hospitals, these were to determine the volume of data to be reported for the SACT reimbursement scheme, the location of the required data within the hospital, as well as establishing contact details of the hospitals’ appointed data managers. The volume of data, used in conjunction with the results of the time study, allowed the researcher to determine the personnel requirement to report these data nationally. The location of the data directed the researcher to hospital computer systems that record these data electronically, thereby having the potential to decrease the personnel burden if the electronic data was reused for the SACT reimbursement scheme. This informed the recipients of the vendor questionnaire.

The main aims of the vendor questionnaire were to verify the systems in use, to establish the potential within the existing systems to record the required data, and to determine what reporting options were available to users. The responses to this questionnaire allowed the researcher to explore the potential to use existing systems to reduce the challenges to the secondary use of these data.

Table 4.4: Research objectives of the questionnaires

Research method - Questionnaires (Section 4.5.2)	
Research aim	**Research objectives**
To investigate the selected case in the areas of the challenges and facilitators to the collection of health data for secondary use.	**1. Hospitals** To determine: • The location of the data required for the SACT reimbursement scheme • The availability of the required data electronically **2. Vendors of selected software system** • To validate systems in place in the hospitals • To determine the potential of existing software systems to: - Record the required data - Report the required data
To investigate the selected case in the area of the personnel resources required to input the data of the selected case.	**1. Hospitals** To determine: • The activity of the hospitals in terms of the number of SACT items prepared and the number of patients treated with SACT as an indicator of the volume of data to be reported • The contact details of persons reporting that data for the purpose of the time study

4.5.3 Quantitative – time study

The aim of the time study, the second component of the quantitative research, was to establish a best estimate of the average time required by the participants to input the required data as described in Table 4.5 below.

Table 4.5: Research objectives of the time study

Research method	**Time study (Section 4.5.3)**
Research aims	**Research objectives**
To investigate the personnel resources required to input the data of the selected case.	Establish the average time required to manually input these data so as to generate an estimate of the time required to input the required data at a national level

This average task time could then be utilised to estimate the time required at a national level to submit these data when combined with the hospital activity identified in the hospital questionnaires, as described in Section 4.5.2. This would demonstrate the personnel impact of the data input of the selected case. Observational techniques, frequently used in time and motion studies, present an opportunity to collect information first hand and have the potential to reveal unanticipated information (Groover, 2007). The time-consuming and labour-intensive nature of these studies are considered disadvantages (Fink, 1993) and resulted in the researcher deciding to limit the number of participants due to their dispersed location and the time constraints of this research. Selection of representative participants, for the purpose of the time study, relied on opportunistic sampling and availability, but intended to include a minimum of four hospital participants familiar with the data entry process of the SACT reimbursement scheme. Consideration must also be given to the Hawthorne Heisenberg phenomenon as those being observed may be performing in a manner that is atypical, resulting in the actions observed not being reflective of the group as a whole (Fink, 1993, Groover, 2007). The researcher attempted to overcome this by incorporating repetitions of each task. A structured data tool was utilised in order to produce dependable data (Fink, 1993) and is included in Appendix 7.

The first three of Groover's procedures for direct time study were utilised as the basis of the experimental design, and the elements of good practice in conducting time studies were utilised in the construction of the research approach (Magagnotti and Spinelli, 2012, Groover, 2007), as described in Table 4.6.

Table 4.6: Components of the time study and the research approach. This table includes the good practice components of time study (Magagnotti and Spinelli, 2012) and the first three of Groover's procedures for direct time study (Groover, 2007)

Time Study Component (Magagnotti and Spinelli, 2012)	**Research Approach**
Research setting	Place of employment of the person being observed
Experimental design	• Define and document the standard collection procedure • Divide the task into work elements • Time the work elements Source: (Groover, 2007)
Time data collection	Utilising a stop watch (Telstar®)
Data analysis	Establish average times for task elements for overall results and for each participant.
Reporting	The time required per annum to report these data based on average times as calculated and volume of data as ascertained from the hospital questionnaires.

The data-reporting process, as described in Chapter 3, was broken into short, relatively homogeneous work elements, as described in Table 4.7. This allowed the breakdown of the timing into short discrete elements, in addition to enabling the researcher to become familiar with the process.

Table 4.7: Work elements of the data input process

Work element of the data-reporting task	
1	Login
2	New patients – i.e. cycle one of chemotherapy a. Patient registration – existing patient on PCRS file – add three patients b. Patient registration – patient not on PCRS file – add three patients
3	Finding a patient – patients already registered – find three patients from above
4	Changing a patient’s diagnosis – change the diagnosis of the three patients from point 3.
5	Adding drugs utilised in the treatment of the patient – add one line of drug data for three patients and two lines of drug data for three patients (may be the same patients)

The time studies were conducted at the place of the person’s employment, utilising the structured data collection tool which clearly identified the start and stop times to ensure validity between measurements. The start time was defined as the user's first orientation towards the application, following any explanation and discussion, and the end time was defined as the point when the user signalled verbally that the task had been completed. The researcher corrected the time taken to complete the tasks in a number of instances where interruptions happened, when users made errors in data entry or undertook an incorrect task. The live system was utilised where possible, and in all other cases a test system was utilised. The mean, median, range and standard deviation was calculated for all timings in addition to the mean time of each participant.

4.5.4 Quantitative – data fields

The researcher compared the data fields of the SACT reimbursement scheme and other national data collections from secondary care in order to establish if variations existed between the datasets, as further described in Section 5.6.

4.6 Conclusion

There are many challenges to the secondary use of data, making it a complex phenomenon suitable for research in a real life context. As concluded in Section 4.3, the researcher selected an embedded single-case design case study strategy for the enquiry, where the embedded subunit research further explored one aspect of the selected case and the qualitative and quantitative data were collected simultaneously. Attention was paid to the case study design to maintain a chain of evidence, remove bias and ensure validity.

A mixture of quantitative and qualitative methods was used for the purpose of this research, including interviews, questionnaires and a time study, in addition to the literature review, to provide direction for procedures in the research design and to answer the research aims (Creswell, 2009, Yin, 2009).

The methodology employed facilitates triangulation and integration of the literature review as well as the research findings in order to describe the challenges and facilitators to the collection of health data for secondary use in the selected case, in addition to informing recommendations beyond the research domain. Chapter 5 and Chapter 6 contain the quantitative and qualitative research findings revealed using the methodology discussed above.

Chapter 5. Quantitative Research Findings

5.1 Introduction

This chapter describes the quantitative findings from the research questionnaires and the time study, integrated with the findings of the literature review, and demonstrates their relevance to the research questions. The key quantitative findings are summarised for integration into the main discussion in Chapter 6.

5.2 Questionnaire Results

As discussed in Section 4.5.2, one component of the quantitative research involved two online self-administered questionnaires, one directed to hospital pharmacists and another to the three vendors of the software system.

5.2.1 Participation

All 26 centers involved in the administration of SACT were invited to participate in the research to establish a national overview. The hospitals were grouped either as specialist centres or linked centres, as detailed in Appendix 13; the hospital questionnaire is included in Appendix 9. Non-respondents were followed up by email and telephone, for a total of three follow-ups. The initial contact and follow-ups were recorded in an Excel® spreadsheet.

Consent to participate in the research was received from 25 of the 26 invited hospitals. The breakdown between specialist centre response rates and those of the linked hospitals is shown in Figure 5.1.

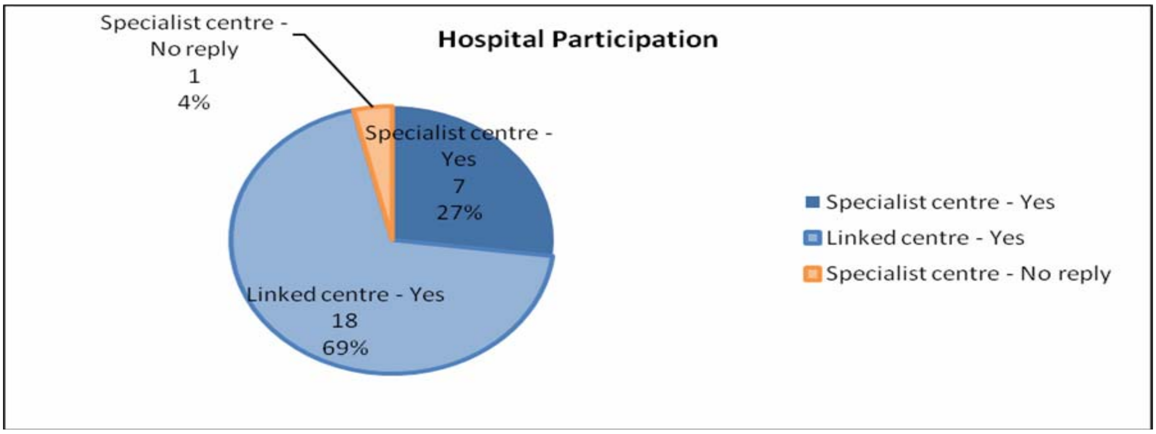


Figure 5.1: Hospital questionnaire participation

The responses to the hospital questionnaire allowed the researcher to establish that the data required for the SACT reimbursement scheme already existed in an electronic format in software systems used to support the preparation of SACT. This is further detailed in Section 5.2.3. There were a total of five different DASCT preparation systems, from three vendors. These vendors were invited to participate in the research, which is included in Appendix 10. Three complete responses were received.

5.2.2 Response rates to the key sections of the hospital questionnaire

The main research objectives addressed by the hospital questionnaire, as described in Section 4.5.2, were to elicit information in four key areas: software systems in use in the delivery of systemic anti-cancer therapy (SACT) in hospitals, which data elements of the SACT reimbursement scheme minimum data set were already collected by the hospitals, the hospital activity in terms of the number of SACT items prepared and numbers of patients treated with SACT, and contact details of the persons inputting the data. The questionnaire sections are detailed in Table 5.1 below, aligned to the dissertation research objectives.

Table 5.1: Hospital questionnaire sections aligned to research objectives

Section		Research objective
3	Details of software systems used in your hospital to support prescribing, preparing, dispensing or administering SACT.	To establish the variety of systems in use
4	Availability of data	To establish the location of the required data in the hospitals.
5	Activity as number of items prepared and numbers of patients treated	To establish the volume of data that would be required to be reported per hospital
6	Contact details of person engaged in web portal data entry	To establish the contact details of the persons inputting the data.

Each section, identified in Table 5.1, had a number of subsections. The completeness of the responses to these sections were classified, for the purpose of this research, as “complete”, “partial” or “no response” as described in Table 5.2.

Table 5.2: Classification of response rates to the four main sections

Section		Complete	Partial	No response
3	Details of software systems used in your hospital for the listed applications	Responses to all 4 subsections, including Yes, No or Unknown	Responses to 1-3 sections, including Yes, No or Unknown	No response
4	Location of data	Minimum of one response to each data line	Response to a minimum of one data line	No response
5	Activity as number of items prepared and numbers of patients treated	Each subsection had a numeric or text answer	At least one subsection had a numeric or text answer	No response
6	Contact details of person reporting the data	Answer = No or Yes (minimum of one contact)	Answer = YES but no contact details given	No response

The response rates to the four key questionnaire sections are shown in Figure 5.2. There were 22 complete responses to the question on hospital systems in use; this contrasts sharply with the responses to the question on the location of the required data within various hospital systems, with only one complete answer. There were also substantial gaps in the information supplied on the hospital activity which included items prepared and numbers of patients treated.

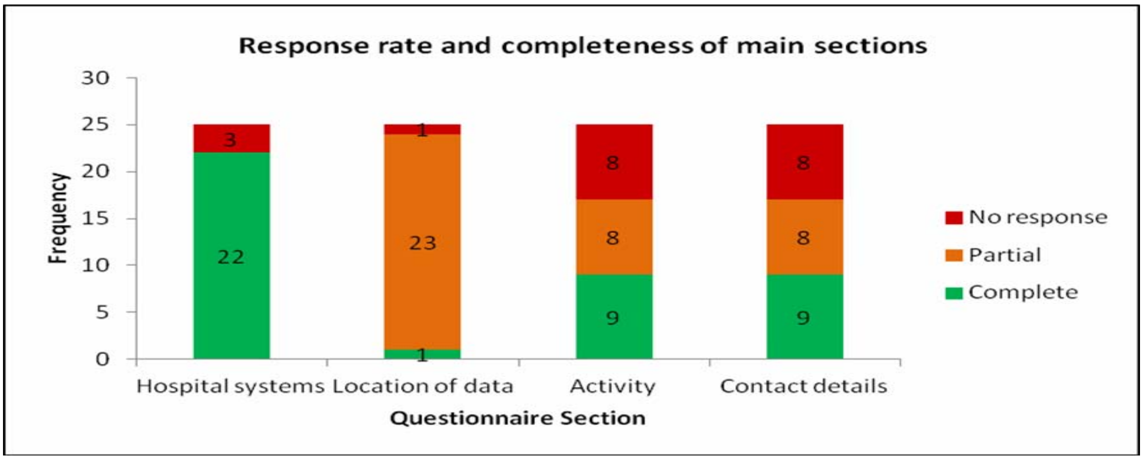


Figure 5.2: Hospital questionnaire response rate to the four main sections of the hospital questionnaire

5.2.3 Systems in use to support SACT processes in hospitals

As mentioned above, there were 22 responses to the questionnaire section on hospital systems in use, revealing disparities in the systems used in the various hospitals, as shown in Figure 5.3. All 22 hospitals had a pharmacy dispensing system and 18 hospitals had both a SACT preparation system and a pharmacy dispensing system. The utilisation of clinical systems to support electronic prescribing systems and electronic medication administration systems was less prevalent.

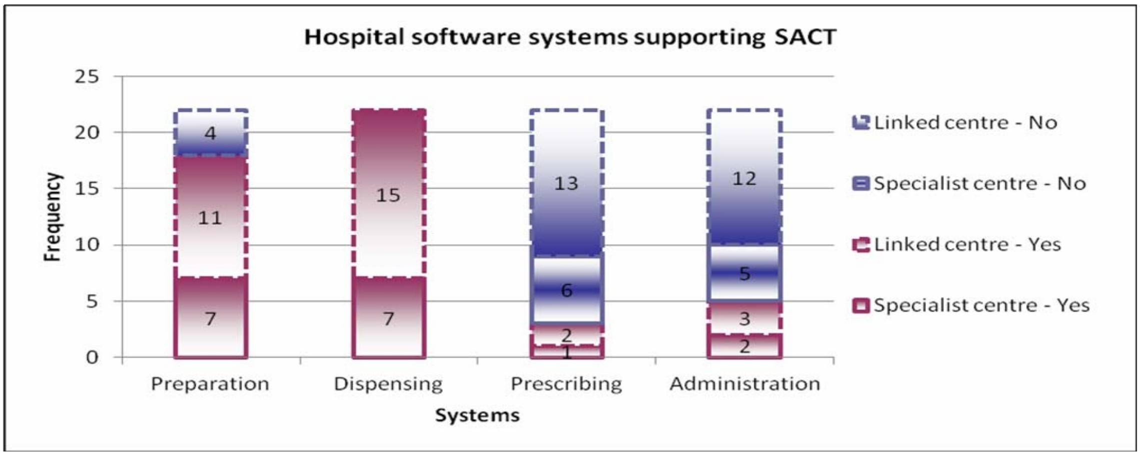


Figure 5.3: ICT systems in use in hospitals to support SACT services **-** Reported utilisation rates of software systems that support SACT prescribing, preparation, dispensing, or administration

5.2.4 Location of the data

The objectives of this section of the research were to ascertain if the storage format of the required data was paper or electronic, in addition to ascertaining if there was a common software system with the potential to record all the data required for the SACT reimbursement scheme.

The responses indicated that the required data was stored in paper charts and paper SACT prescriptions, as well as in various software systems such as hospital information systems (HIS) and SACT prescribing, preparation, dispensing or administration systems. The questionnaire responses, as displayed in Figure 5.4, showed that the demographic data were widely available; both in paper format and in many of the hospitals' software systems, with the exception of the PPSN. The research data also showed a deficit of the other SACT reimbursement scheme patient identifiers recorded in the hospital systems. The researcher had sought information on the availability of a hospital medical record number (MRN), which was widely available. This demonstrated the challenges presented when systems designed for use in one domain, primary care, are utilised in a different domain, secondary care; where both domains commonly utilise different patient identifiers.

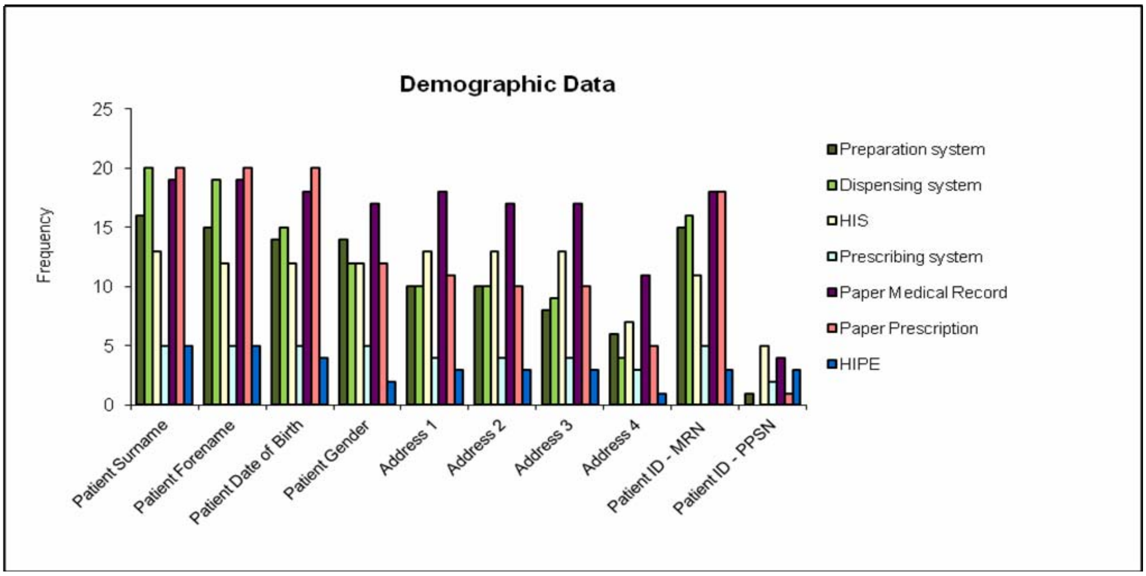


Figure 5.4: Stored locations of demographic data

The results found that clinical and drug data were mainly recorded on paper, with less of these data recorded electronically in various software systems, as can be seen from Figure 5.5.

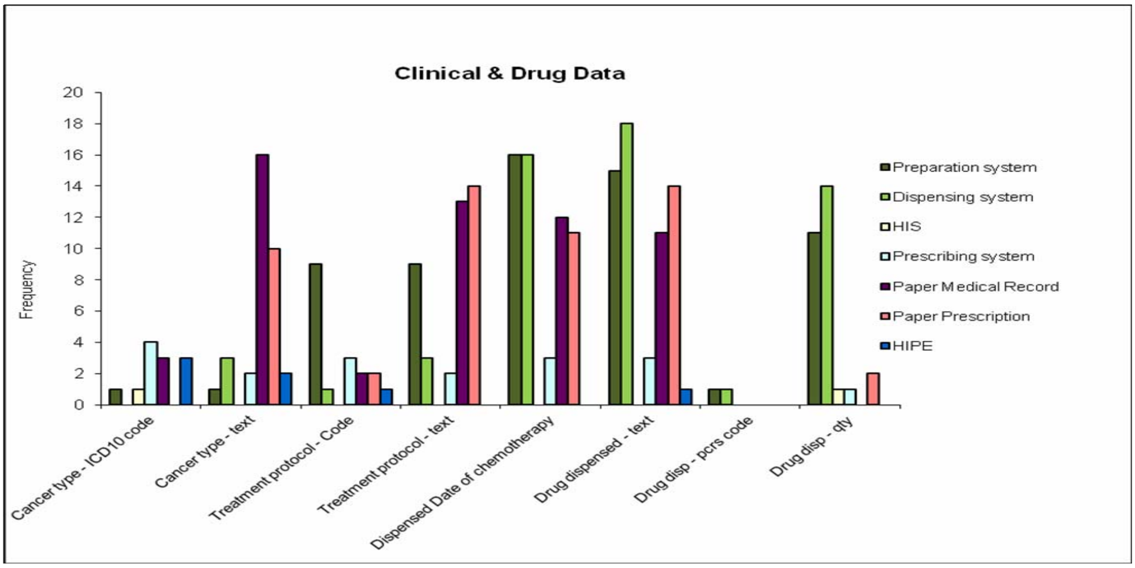


Figure 5.5: Stored locations of clinical and drug data

The pharmacy SACT preparation systems and the pharmacy dispensing systems were the most commonly used software systems reported to contain the required clinical and demographic data. The ICD10 code was absent from the dispensing systems and only one response confirmed the presence of the PCRS drug code. This contrasted with the wide availability of the local patient identifier, the MRN. The prescribing systems also recorded the required data, but were in use in only three hospitals.

The researcher chose to further explore the pharmacy SACT preparation systems as the potential source of electronic data for secondary use, due to their prevalence and their potential to contain the required data. The responses to the hospital questionnaire allowed the researcher to establish that there were five different SACT preparation systems in use, from three vendors, with a number of linked hospitals not having a SACT preparation system. The hospital responses indicated that 4 hospitals had no SACT preparation system, but the vendor responses showed that only 2 hospitals had no such system. This highlights the disadvantages of self-administered questionnaires, as detailed in Section 4.5.2, as the discrepancy could potentially be due to a misunderstanding or alternatively an error. The data from both sets of questionnaires were combined, which resulted in information for all 26 hospitals as detailed in Figure 5.6.

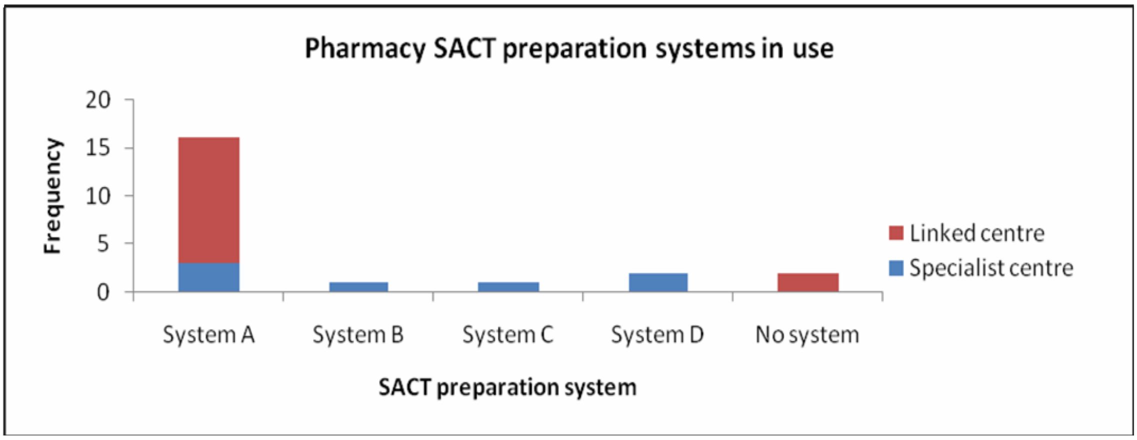


Figure 5.6: Hospital SACT preparation systems in use

The responses to the hospital questionnaire showed that 20 of the 22 hospitals utilised one system, system A, and an additional 2 hospitals used an upgraded version of this system, system D. There was variation in the 16 responses from these hospitals to the question on the location of data within the SACT preparation system. The researcher then combined the hospital questionnaire responses on the location of data within systems A and D, with the vendor's response to examine these variances. The results are shown in Figure 5.7 and Figure 5.8. The system vendor responses indicated that there were data fields available in the current system to record the required data, yet interestingly the hospital responses did not reflect this fact. This led the researcher to conclude that the hospitals chose not to utilise these data fields as they were not required for their processes, or that they were unaware of their existence, or alternatively that the vendor may have misinterpreted some of the fields or the question. This further illustrates the disadvantages of a self-administered questionnaire as a data collection tool, as discussed in Section 4.5.2.

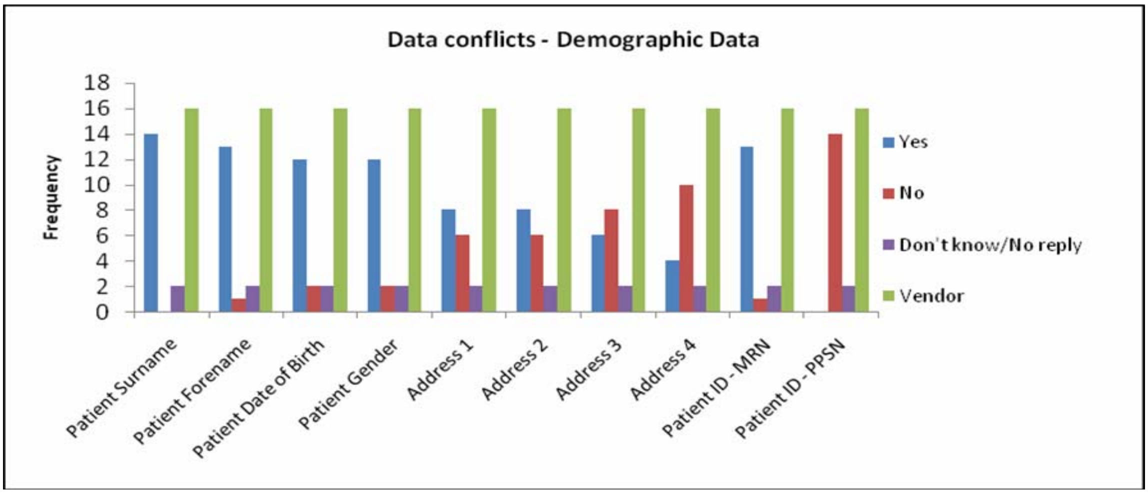


Figure 5.7: Data conflicts - demographic data

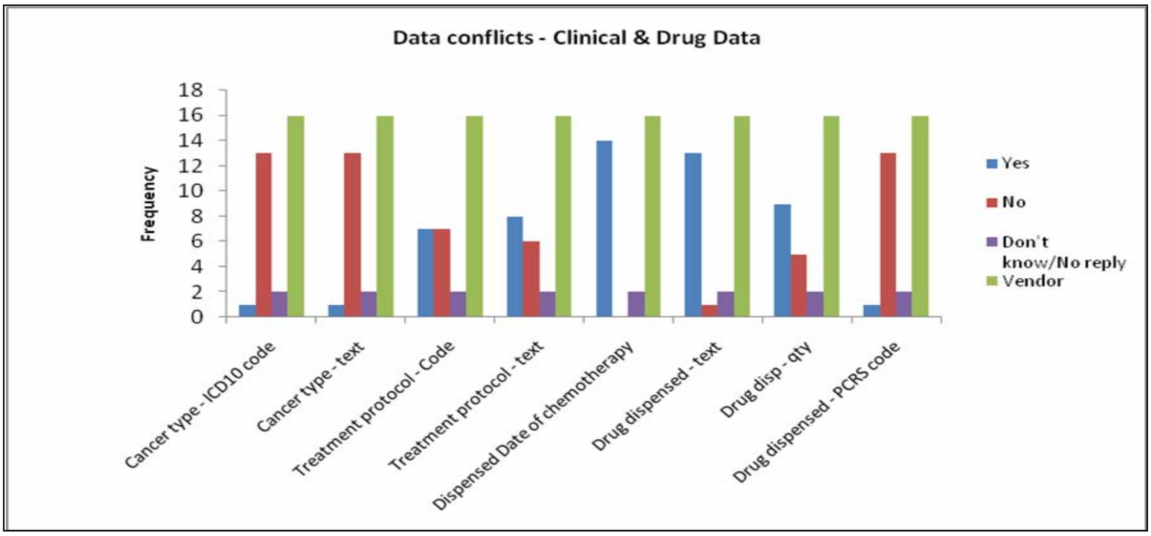


Figure 5.8: Data conflicts - clinical and drug data

5.2.5 Hospital activity

The hospital activity section of the questionnaire sought information on the number of SACT items prepared and numbers of patients treated with SACT. The responses were graded, as described in Table 5.2, as complete, partial or no response. There were mixed grades of response for the SACT preparation system A and those with no system, as displayed in Figure 5.9.

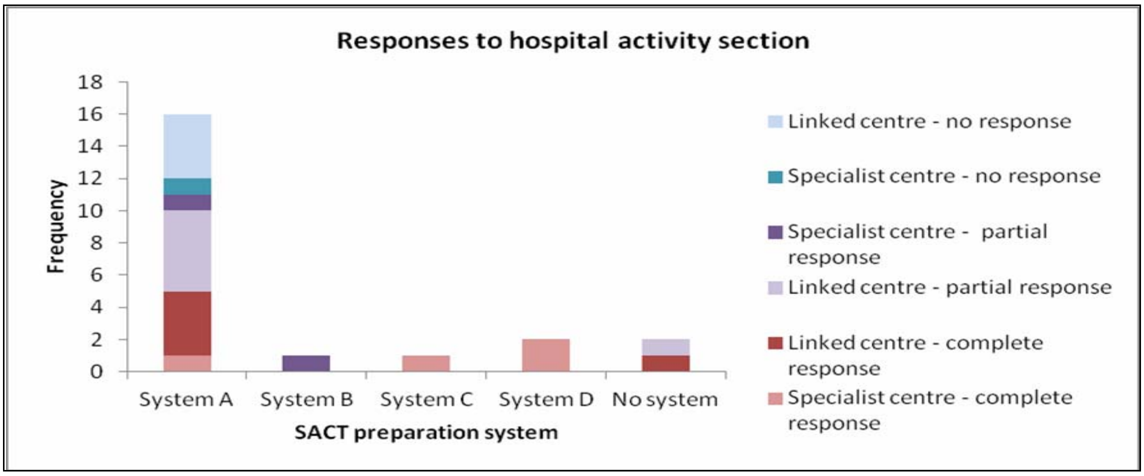


Figure 5.9: Hospital activity responses graded as complete, partial or no response

The researcher had invited all 26 hospitals to participate in this research with the intention of compiling a national picture of hospital activity. Due to the variation in the level of response to this section of the questionnaire, additional data was requested from the NCCP gatekeeper in order to complete the dataset. These data had been independently compiled by the NCCP subsequent to the commencement of this research.

The responses to the hospital activity section, as shown in Appendix 12, are anonymised, with the NCCP data differentiated from the hospital questionnaire data. The availability of the independent data set allowed the researcher to explore the quality of the reported data. The numerical responses were equal in most cases, with some minor variations. The response rates to the questionnaire and the NCCP responses are detailed in Figure 5.10, neither of which contained responses from all the hospitals. The combination of both data sets resulted in responses being available from 24 hospitals. The collection of these data was essential to allow the researcher to investigate the personnel resources associated with the entry of these data as a subunit of the selected case.

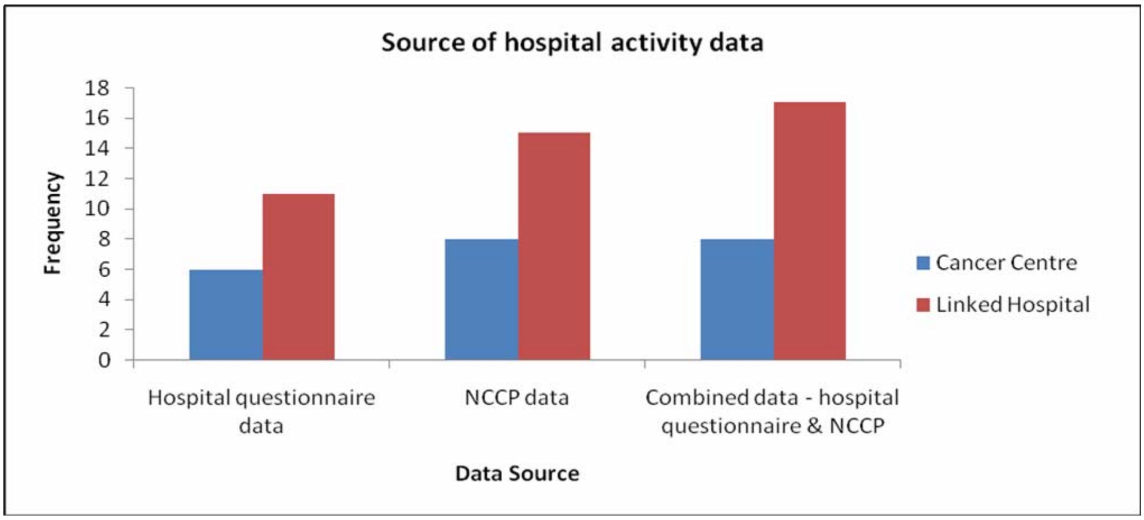


Figure 5.10: Hospital activity data by source - NCCP or hospital questionnaire

The researcher had anticipated some variation in the responses to this section of the hospital questionnaire, as at an early stage of the research a number of queries had been received from participants concerning difficulties being encountered in the extraction of the data for the questionnaire. Consequently, the researcher advised these participants on the extraction of comma-separated value (CSV) files containing the required data. Some of the responses to the subsections of the hospital activity section described difficulties in data extraction, rather than the required data. Samples of these responses are detailed in Table 5.3. This is of interest as some hospitals utilising the same SACT preparation system were capable of extracting the required data and others were not - an outcome which could be attributed to differences in user training and expertise in system use, as discussed in Section 2.8.3.

Table 5.3: Selected free text comments on the difficulties of data extraction. These comments reveal that some participants encountered difficulties in the extraction of hospital activity data.

Question	Comment	System
Number of existing patients	Impossible to quantify. System A is not capable of running this report. Not searchable with current software. Information not readily available.	System A
	Data not readily available.	No SACT preparation system – dispensing system only.
Number of patients on specified drugs	System A is not capable of running this report. Not searchable with current software.	System A
	Data not readily available	No SACT preparation system – dispensing system only.
Patient numbers generally	HIPE data is available on total attendances but not new patients, as HIPE is based on episodic reporting	Various
Number of doses of specified drugs	Information not readily available.	System B
	Not readily exactable System A is not capable of running this report Information not readily available	System A

The hospital activity had been requested for the 2012 calendar year, and is summarised in Table 5.4 below, by cancer centre and linked hospital. The descriptions of the data items requested are included in the questionnaire in Appendix 9, and include numbers of SACT items prepared during the year; the total numbers of patients receiving SACT, as well as those receiving a subset of SACT. As mentioned in Section 5.2.5, there was variation in the numbers of hospitals responding to the questions on hospital activity; this can be seen from Table 5.4, where N indicates the number of replies from cancer centres or linked hospitals. A national hospital activity annual total was then estimated, based on the mean of the responses to each question and the total number of hospitals involved in SACT services, this being 8 cancer centers and 16 linked centers respectively.

Table 5.4: Estimated annual national hospital activity data

		N	Mean	Median	Min	Max	Range	Estimated annual totals
Patients numbers – total	Cancer centre	8	888	880	599	1,284	685	7,105
Patients numbers – total	Linked hospital	13	238	215	38	571	533	4,292
Patients numbers - specified SACT	Cancer centre	8	211	183	153	289	136	1,685
Patients numbers - specified SACT	Linked hospital	15	52	36	6	178	172	941
New patients	Cancer centre	7	506	435	362	797	435	4,046
New patients	Linked hospital	12	154	128	25	371	346	2,775
Total items	Cancer centre	8	11,662	10,373	8,166	20,084	11,918	93,299
Total items	Linked hospital	16	3,883	2,679	548	10,278	9,730	69,893
Total items specified SACT	Cancer centre	6	1,995	1,774	1,296	2,848	1,552	15,963
Total items specified SACT	Linked hospital	15	606	208	4	2,473	2,469	10,900

5.2.6 Contact details of data managers

Contact details of the hospital data manager submitting the SACT reimbursement scheme data were requested in the hospital questionnaire to enable the researcher to invite them to participate in the time study. 25 hospitals responded to the hospital questionnaire and of these 17 completed this question. 8 responses indicated that no data manager had been nominated and the remaining 9 provided a minimum of one contact name each, resulting in a total of fifteen data manager contacts being established.

5.3 Time Study Results

This component of the research was important to determine the personnel resources required for the input of the required data. The time studies were undertaken to determine the time required to enter the SACT reimbursement scheme data, and were conducted at the place of each participant’s employment, using the data collection tool detailed in Appendix 7. The task of the data input was broken into five short, relatively homogeneous work elements, as described in Section 4.5.3, and detailed in Table 5.5 below. Where possible, participants were asked to repeat each task a minimum of three times to overcome the Hawthorn Heisenberg phenomenon, as described in Section 4.5.3, and to ensure that the time study would not require more than thirty minutes of each participant’s time, as determined when piloting the time study data collection tool.

The time studies commenced with a number of logins to allow the participants to become familiar with being observed.

Table 5.5: Data input work elements

	Work element
1.	Login
2.	Patient registration a. Existing patient on PCRS file (only available on live system) b. Patient registration – patient not on PCRS file
3.	Finding a patient – patients already registered
4.	Changing a patient's diagnosis
5.	Adding drugs utilised in the treatment of the patient - add one line of drug data for a patient

5.3.1 Sample observed

The number of participants observed was limited due to the time constraints of the research and the dispersed locations of the persons, as discussed in Section 4.5.3, in addition to the small contact list of 15 people generated from the questionnaire. Due to these limitations the researcher also invited a number of NCCP staff involved in the implementation and support of the SACT reimbursement scheme to participate, resulting in a total of 9 participants being observed.

Details of the participants, including their experience of the patient registration for the SACT reimbursement scheme, are shown below in Figure 5.11. All but one of the participants utilised the test environment. This resulted in very limited data, that of a single user, to be collected on work element 2(a) - registration of existing PCRS patients to the SACT reimbursement scheme, as detailed in Table 5.5 above.

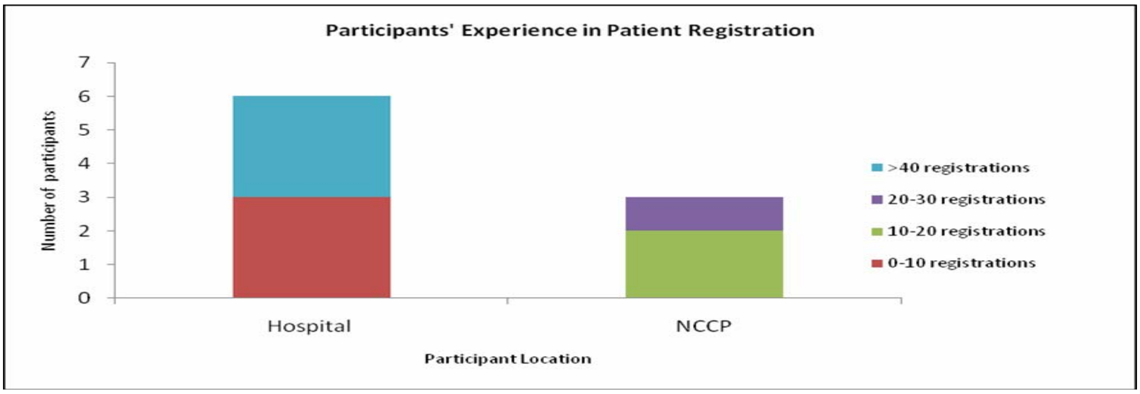


Figure 5.11: Participants' experience in patient registration

5.3.2 Time study results

Table 5.6, generated from the raw data, summarises the data collected in the form of various statistical functions, broken down by work element. The results of each work element have been calculated two ways, one using all the recorded times, where the number of occurrences (N) is the total number of times recorded for that work element, and the second using the mean time of each participant, where the number of occurrences (N) is the number of persons observed for that work element. The mean of all timings differed from that of the user means wherever the number of times recorded per participant varied.

Table 5.6: Time study results (Seconds)

		N	**Min**	**Max**	**Mean**	**Median**	**Range**	**σ**
Login	All timings	29	14.80	30.00	20.51	19.70	15.20	3.912
	User Means	9	18.38	25.83	20.64	19.28	7.46	2.630
Add new patient from PCRS file	All timings	3	69.00	101.70	81.90	75.00	32.70	14.213
	User Means*	1	81.90	81.90	81.90	81.90	0.00	0.000
Add new patient (not on PCRS file)	All timings	27	90.80	167.00	116.07	108.90	76.20	19.327
	User Means	9	99.13	151.00	116.07	111.13	51.87	16.402
Select existing patient	All timings	31	17.20	44.80	30.21	29.10	27.60	6.754
	User Means	9	22.93	39.32	30.52	28.65	16.39	5.313
Patient verification	All timings	25	7.50	13.20	10.11	10.00	5.70	1.369
	User Means	6	9.70	10.60	10.06	10.01	0.90	0.296
Change patient diagnosis	All timings	26	12.70	36.90	22.48	21.30	24.20	5.261
	User Means	9	16.70	31.85	22.80	21.13	15.15	4.427
Add drug line	All timings	38	14.90	35.80	25.36	25.45	20.90	4.603
	User Means	9	20.18	28.83	25.51	26.17	8.65	2.903

*Note – only 1 user observed using the live SACT reimbursement scheme

The researcher then combined the different work elements, as detailed in Table 5.7, to form the three discrete work processes associated with the data entry into the web portal, as described in Chapter 3.

Table 5.7: Web portal work processes

	Login	**Add patient to system**	**Select patient**	**Verify patient**	**Change diagnosis**	**Add drug line**
Register a patient	Y	Y		Y		
Change a patient's diagnosis	Y		Y	Y	Y	
Register a claim	Y		Y	Y		Y

5.4 Personnel Resources Required for Data Input

One of the research questions aimed to investigate the personnel resources required per annum to manually input the SACT reimbursement scheme data. The researcher made a number of assumptions in the estimation of the personnel resources, as detailed in Table 5.8. These assumptions were based on an expected efficiency, where the researcher pragmatically expected that each data manager login would result in the entry of more than one data line or the registration of more than one patient. The assumptions included a single registration of patients, a maximum of 5% of patients requiring a diagnosis change, and 80% of patients being preregistered on the PCRS database. The limitations of these assumptions include underestimation of the personnel resources in those cases where the assumptions as outlined are not met. This is particularly true in the case of linked hospitals with low levels of activity, as displayed in Table 5.4, where the number of new patients per annum is as low as 25 and the minimum total SACT items per annum is as low as 548. In this case, the login events required would be more frequent and would include fewer data lines than included in the assumptions in Table 5.8 below, resulting in an increased personnel requirement.

Table 5.8: Assumptions for personnel resource calculations

	Assumption
1.	There would be one person nominated as a data manager in each hospital
2.	The hospital data manager would complete an average of • 50 claims per login episode or • 10 patient registrations per login episode
3.	Each patient would be registered only once and not require registration in multiple hospitals
4.	No more than 5% of new patients registered would require a diagnosis change
5.	The demographics of 80% of the patient population to be registered to the SACT reimbursement scheme already exist on the PCRS database

The total time required per annum was then calculated using the data from Table 5.4 and Table 5.6 and based on the scenarios in Table 5.9.

Table 5.9: Scenarios utilised to estimate the data input time per annum

Scenario	**Summary**	**Comment**
1	Register all new patients on all drugs and claim for those drugs	Uses the total number of drugs administered per year and total number of new patients per year
2	Register only new patients on specified drugs and claim for those drugs	Uses an estimate of new patients per annum on specified drugs based on the percentage of patients administered those drugs in 2012 and the total numbers of the specified drugs
3	Register all existing patients on all drugs and claim for those drugs	Uses the total number of drugs administered per year and total number of patients per year
4	Register all existing patients on specified drugs and claim for those drugs	Uses patient numbers on specified drugs and the total numbers of the specified drugs

The estimated time required to input the data per annum is shown in Table 5.10. This time has been converted to hours required and calculated for each scenario using the mean of all the timing in addition to the mean time of each participant, as discussed in Section 5.3.2.

Table 5.10: Estimated time required per annum for data input (Hours)

Scenario		Min	Max	Mean
Register all new patients on all drugs and claims for those drugs	All timings	1,968	4,588	3,228
	User Means	2,599	3,851	3,247
Register only new patients on specified drugs and claims for those drugs	All timings	341	787	556
	User Means	448	661	559
Register all existing patients on all drugs and claims for those drugs	All timings	2,059	4,747	3,358
	User Means	2,711	3,975	3,377
Register all existing patients on specified drugs and claims for those drugs	All timings	348	825	583
	User Means	472	692	587

There were no personnel resources allocated to this data collection (O'Reilly, 2013). The literature review, as discussed in Section 2.8.1.3, confirms that data collection places an increased burden on the providing institution. The results of the time study, combined with the hospital activity, in terms of the number of SACT items prepared and the numbers of patients treated with SACT, as collected from the quantitative questionnaire, and shown in Table 5.10 above, clearly demonstrate the burden of this data collection. Utilising a 39 hour week as an average week, the mean time per annum required to register all new patients on all drugs, which is the data required ultimately by the NCCP, is 3,228 personnel hours, equating to 83 personnel weeks. There would be a similar initial personnel requirement to register all those patients on current treatment. Smaller scale data collections, focussing only on patients on a subset of SACTs would require less data input, but will only supply a subset of the data required.

The personnel burden as calculated above is likely to be underestimated and the real resource implications would tend towards the maximum range of the scale, resulting in a personnel burden of 4,588 personnel hours per annum, equating to 118 weeks. There would be an equivalent requirement to register all existing patients and their drugs. The rationale for selecting the maximum time is based on a number of reasons; the most important being that all bar one of the time studies was conducted using the test system.

The researcher observed that the test system environment did not encourage the participants to quality-check their data, whereas in contrast, the one participant utilising the live system was cautious in patient selection for registration, generating claims and crosschecked the data. In addition, the low volumes of data reported in some of the linked hospitals may result in logins which do not achieve the efficiencies of a minimum of 10 patient registrations or the entry of 50 data lines. This will result in a higher impact for data entry in these sites.

It should also be borne in mind that the personnel burden as detailed above pertains only to the time required to report the data and does not include the time required to collate these data and to manage the associated reimbursement transactions.

5.5 System Reporting Ability

The reporting ability of the hospital SACT preparation systems had formed a component of the vendor questionnaire in order to ascertain if the data required for the SACT reimbursement scheme was available as a report that would facilitate the extraction of the required data from the system. The researcher also wished to establish if the systems' reporting capabilities and export format contributed to the difficulties encountered by the SACT preparation system users in extracting the data for the purpose of this research, as detailed in Section 5.2.5. The report export formats are detailed, in Table 5.11 below. The capability of each system to allow generation of user-defined reports (UDRs) from the systems' raw data is also detailed.

Table 5.11: Vendors' responses to report export format

Report type	System A	System B	System C	System D
User designed reports (UDR)	No	No	Yes	Yes
Report export – excel	No	Yes	Yes	No
Report export – PDF	No	Yes	Yes	No
Report export – CSV	No	Yes	Yes	No
Report export – CSV	Yes	Yes	Yes	Yes
Report export – XML	No	Yes	No	No

The questions with regard to the availability of specific data as a standard report (SR), a user designed report (UDR) or potentially as a new standard report (NSR), were informed by the data required for the purpose of this research. The objective was to ascertain the ease with which these data could be extracted, and therefore the systems potential to supply the required data.

The system vendor's reply indicted that the majority of these data could be extracted as standard reports, with the exception of system A, which did not have a standard report available on the number of patients receiving a specified list of drugs. This information, in combination with the difficulties encountered by some participants in extracting the data, as discussed in Section 5.2.5 and Table 5.3, led the researcher to conclude that there was more benefit to be garnered from the current systems through additional training and support of system users in the use and capabilities of the systems. This is also supported by the literature, as discussed in Sections 2.6.1 and 2.8.3, where training and education, in addition to process change, are required to realise the full potential of implemented systems.

Table 5.12: Vendors' responses to available reports

Reports available	**System A**	**System B**	**System C**	**System D**
Patient numbers – dispensing of SACT	SR	SR	SR	SR
Number of patients who receive specific drugs		SR	SR/UDR/NSR	SR
Number of items prepared	SR	SR	SR/UDR/NSR	SR
Number of items of specific drugs prepared	SR	SR	SR/UDR/NSR	SR
By generic name	Yes	Yes	Yes	Yes

5.6 Data Fields of the Selected Case

The data required in the selected case, as detailed in Appendix 4 and Appendix 5, were compared to other national data collections in the domain of cancer, as discussed in Section 4.5.4. Some of the data fields required for the SACT reimbursement scheme are also collected in secondary care national data collections, as shown in Appendix 5. There are variations between the data dictionaries utilised in the different data collections, an example of this being the variations in the "County" fields, where each data collection utilises a different data dictionary, as detailed in Appendix 6.

It is worth noting that the PCRS ordinarily have little involvement in the funding of drugs in secondary care (HIQA, 2013a), which in part explains the absence of the required primary care identifiers in the hospital systems in use.

5.7 Key Quantitative Findings

The quantitative section of this research has shown that, aside from the PPSN and other primary care patient identifiers, the data required for the SACT reimbursement scheme are currently being collected in hospitals; both on paper and also electronically in disparate software systems. The deficit of required patient identifiers may be attributed to the utilisation of primary care identifiers in a secondary care data collection from secondary care, which utilises local identifiers such as the MRN. This emphasises the requirement for a unique health identifier utilised across all health domains in Ireland. Hospitals may also be reluctant to collect the PPSN due to the Data Protection Commissioner's position, that the PPSN should not be used as a health identifier, as discussed in Section 2.8.4.1, in addition to the requirement to inform patients, and receive their consent, to utilise data collected for their clinical treatment for the purpose of reimbursement.

This research has also identified some variations in the datasets of national data collections, emphasising the need for cooperation in the development of national data collections and reuse of datasets, as discussed in Sections 2.8.1.1 and 2.8.4.3.

The SACT preparation systems used by 24 of the 26 hospitals have the potential to record all the required data, but are not currently utilised to do so in the majority of the sites. A work process change could allow the recording of the required data electronically, for potential output to the SACT reimbursement scheme or, at a minimum, to streamline the data collation internally for manual reporting. A thorough investigation into the availability of data, in advance of the implementation of the SACT reimbursement scheme, may have revealed this potential, as discussed in Section 2.8.1.3, and thereby may have led to an alternative, less resource-intensive method for this data collection process through the reuse of electronic data. This also highlights the requirement for national and local leaders to ensure that ICT projects are implemented fully and aligned to national objectives, as discussed in Section 2.8.3.

This component of the research revealed difficulties experienced by the participants during data extraction from SACT preparation systems. Additional training could improve data extraction in addition to facilitating a more complete utilisation of these systems, as discussed in Sections 2.8.3 and 5.5, resulting in a decrease of the data-reporting burden and associated personnel requirements.

This research identified a personnel requirement of 118 personnel weeks per annum nationally, more than two full time personnel, to input the required data for the SACT drugs for all patients. An equivalent resource would be required to register existing patients and their drugs. This is a challenge in the current economic environment of recruitment moratoriums and decreasing numbers of staff, as discussed in Section 2.8.5. The personnel requirement, as identified above, pertains only to the manual input of the data in the selected case. It does not include the time taken currently to collate these data from paper and electronic systems; nor the time required to reconcile the financial statements generated by the reimbursement claims; nor does it address the anticipated growth in numbers due to the rising incidence of cancer and the increasing number of treatments, as discussed in the literature review in Chapter 2. If the reporting of these data is to be fully implemented and sustained, it will require re-evaluation of the method of collection and implementation of facilitators to improve the process.

5.8 Conclusion

The quantitative component of this research has answered the research question posed in Chapter 1, as to the personnel resources required to input the data of the selected case. It also revealed some of the contemporary challenges to the collection of health data for secondary use, these being the disparate computer systems in use in the 26 hospitals involved, in addition to the need for increased user training to maximise value of current software systems, and a requirement for ownership of system implementation at both national and local levels. These conclusions will be further discussed in conjunction with the qualitative findings in Chapter 6.

Chapter 6. Qualitative Research Findings & Discussion

6.1 Introduction

The first section of this chapter describes the steps taken to analyse the narrative data of the semi-structured interviews which were conducted to explore the challenges to the collection of health data for secondary use. The findings of the thematic analysis are then presented and discussed, integrated with the quantitative findings as presented in Chapter 5. This integration and triangulation of data sources enhances the data quality, based on the principles of idea convergence and the confirmation of findings as guided by the interpretation of the researcher (Baxter and Jack, 2008).

6.2 Semi-structured Interviews - Data Analysis

A total of 18 stakeholders were interviewed, from March to May 2013, where possible face-to-face and otherwise by telephone, as shown in Figure 6.1. The primary stakeholders in the SACT reimbursement scheme included staff of the NCCP, the PCRS and the hospitals involved, as discussed in Section 2.9, and secondary stakeholders included representatives of HSE ICT, HIQA, and SACT preparation system vendors, amongst others. Additional secondary stakeholders emerged during the course of the research. A complete list of stakeholders and those interviewed is included in Appendix 16.

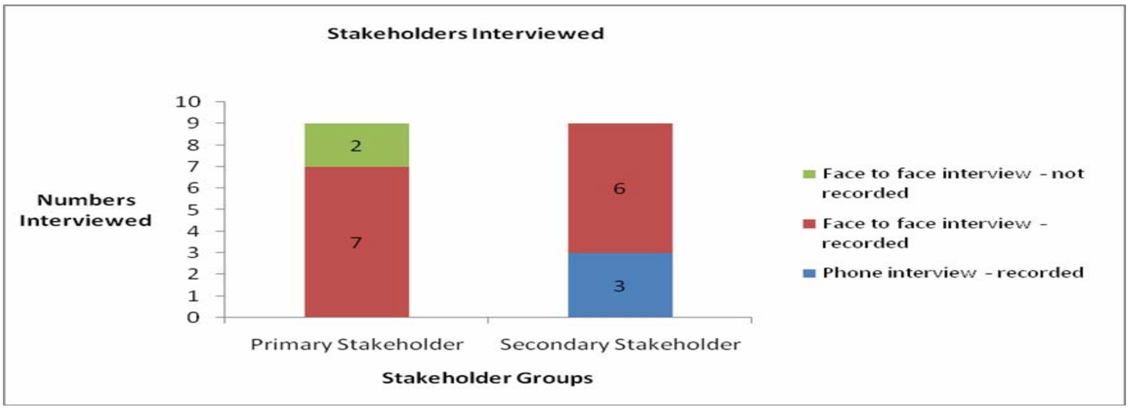


Figure 6.1: Stakeholders interviewed

The interview topics and associated data collection tools were based on the AHRQ framework of the six challenges of today's performance measurement data collection and reporting environment, as described in Table 2.5, in addition to exploring the interviewees' awareness of both the motivation to collect data for public health and cancer control purposes, and the standards used in the domain of health care data collection.

Notes were taken at all 18 interviews, with 16 interviews recorded. The researcher then created MS Word® documents for each interview, based on the interview notes and the interview recordings, using the interview topics as a high-level thematic framework to focus the analysis. Each interviewee received their individual document for verification, with the option of providing an addendum if desired, in order to ensure validity, remove bias and maintain a chain of evidence. This initial analysis allowed the researcher to focus on the topic-specific stakeholder responses of both individuals and groups. The anonymised interview quotations were annotated, when included in the dissertation, to maintain the chain of evidence in the format of Interviewee X: Primary Stakeholder (PS) or Secondary Stakeholder (SS); I2:PS, referring to interviewee 2: primary stakeholder.

Data analysis was manual, using MS Excel® spreadsheets, as the data set was considered small enough to be manageable. The data was initially categorised into high-level themes as described above and then copied into spreadsheet rows; each column was dedicated to a unique interview. Data was replicated where it was found to be consistent with more than one theme. Subcategories identified through the literature review were utilised as preset categories and supplemented by additional subcategories as they emerged, to allow for greater discrimination and differentiation of the analysis. No additional themes emerged. The initial list of categories and subcategories expanded as the data was processed. The addition of new subcategories prompted an iterative revision of previously categorised data. This process continued until all data had been documented and analysed to the point where no new subcategories were emerging. The subcategories were then examined and subcategories duplicated between themes were merged under one theme, reducing the number of subcategories from 156 to 117. The resulting category outline is shown in Table 6.1, with examples of some of the associated subcategories. The final list of categories and subcategories is detailed as a component of the coded thematic analysis extract in Appendix 8.

Table 6.1: Categorisation of interview topics

Interview topic		**Sub categorisation – examples**
Data collection/use in cancer control		
1	Main benefits of collecting data for cancer control	Access to service, improve outcomes, statistics, incidence.
2	Agencies that collect data on cancer	No subcategories applied.
3	Mandatory or voluntary reporting	No subcategories applied.
4	Standards used in the current data collection for cancer control	No subcategories applied.
Challenges to the Secondary use of health data		
5a	Inefficiency	Incomplete data sets, duplicate data entry, manual data entry, no quality data checks, dispersed location of data.
5b	Variations in measurement systems	New or different data requirements.
5c	Organisational and cultural issues	New reporting requirement, work practice change.
5d	Technological barriers for electronic health records	Disparate/legacy systems, lack of legislation, use of standards/identifiers, system interoperability and data sets, data ownership, security and privacy issues.
5e	Economic pressures	No additional resources for staffing, ICT.
5f	Competing priorities	Data protection, patient-specific information to be communicated to third parties, local vs. National priorities, clinical vs. administrative priorities.
Systemic Anti-Cancer Therapy (SACT) – the selected case		
6	Reason for SACT reimbursement scheme	Budget control, data, audit.
7	Stakeholders involved in this scheme	Primary stakeholders, secondary stakeholders.
8	Other data that could be captured which might be of potential benefit	No subcategories applied.
9	Wider role for the data	Financial, national systems.
10	Approach to missing data	Local interfaces, system upgrades, dual approach.
11	Main challenges to the provision of SACT reimbursement scheme data	Subcategorised as question 5.
12	Facilitators to the provision of SACT reimbursement scheme data	Good systems / automated data capture, standard data sets / data dictionaries, accountable lead.
Steps to improve data collection nationally		
13	Steps to improve the national secondary use of data	Subcategorised as question 12.

The findings of the thematic analysis are presented and discussed in the following sections, integrated with the quantitative findings, as presented in Chapter 5. The researcher utilised the AHRQ challenges, identified in the literature review and presented in Table 6.1 above, as a framework, with the exception of data ownership, data protection and privacy, which are presented in a separate section due to their complexity. This allowed collation of the individual intricacies of each challenge, and subsequent presentation of the research findings in a concise and comprehensive way, minimising duplication and avoiding a lengthy narrative, which is a weakness of case study methodology (Creswell, 2009, Yin, 2009). Quotations from the research interviews are included, where relevant, to emphasise the views of the interviewees.

6.3 Challenge - Inefficiency

National cancer information systems (NCIS) are utilised in many countries to centrally aggregate the data required for the implementation of cancer control programmes, as discussed in Section 2.4 of the literature review. The lack of such a system in Ireland generates the requirement for multiple data collections within the domain of cancer control. The interviewees were of the general opinion that the data collected for the SACT reimbursement scheme had other potential applications in the areas of research and cancer intelligence, particularly if aggregated with existing SACT data such as HIPE and the NCRI data sets, to create a NCIS. Additionally a NCIS, receiving automated data feeds from electronic records, would have the potential to remove some of the reporting burden locally, as third parties could run their own reports and analyses and feed these back to local level.

The findings of this research show, that in the domain of cancer, the collection of multiple datasets is being managed by different individuals at local and national level, with disparate data managers appointed to individual collections. This is resulting in the repeated capture of overlapping data fields. The coordination of the collection of these data, either at national or local level, or both, would result in greater efficiencies and decreased cost, as discussed in Section 2.8.1 of the literature review. The reporting methodologies of the data collections also differ, with the SACT reimbursement scheme data and HIPE data submitted to a web portal, whereas other data collections are reported in spreadsheets or word documents, decreasing the likelihood of sharing the data collection. Some data were unavailable for reuse due to their collection in bespoke systems, designed for specific or limited uses, which are not scalable or capable of meeting changing requirements, as discussed in Section 2.6 of the literature review.

Box 6.1: Interview quotations on inefficiency and duplication of work

"There is a lot of time wasted in re-keying data which ends in more data silos e.g. data managers on each site creating individual databases" (I8:SS).

"An army of people (coders) that deal with HIPE data, and another army of people dealing with cancer, another army that deal with this or the other...an awful lot of duplication and fragmentation, variable quality. Not only is this inefficient and results in increased cost but also mitigates against reusing the data and maximising the reusability of the data" (I9:SS).

Efficiency in data collection is also challenged by the fragmented storage of the required data, as discussed in Section 2.5 of the literature review, and further confirmed in the quantitative component of the research, which found that the required SACT reimbursement scheme data was available in paper records, with some data replicated, to varying degrees, in a variety of electronic systems, as discussed in Section 5.2.4. This creates a situation where no single system has all the required data resulting in manual retrieval, potentially from multiple places, to achieve 100% data collection. The interviewees also broadly agreed that, while clinical data is primarily located in paper charts, it may be in more than one paper chart, or stored electronically in various clinical or administrative electronic systems.

The literature review showed that technology, in the form of EHRs, has been shown to improve the quality and efficiency of data capture for both primary and secondary uses of these data (Ottosen, 2012, Sidorov, 2006, Miller et al., 2005, Agrawal, 2002). These systems can collect data as a by-product of patient care rather than as a separate process, supporting the “collect once, use often” paradigm, as discussed in Section 2.6 of the literature review. The findings of this research indicate that data collection for the purpose of the SACT reimbursement scheme is regarded as a process separate from that of data collection in the clinical care of the patient. Interviewees suggested that a different approach, utilising ICT and EHRs, should be considered to improve efficiency and to maximise the use of electronic data already collected in disparate systems.

Box 6.2: Interview quotations on the role of ICT in data collection

"There is a need to optimise automatic capture of data from point of origin, the concept of the single point of data entry and multiple uses of data ” (O'Reilly, 2013).

"We (the HSE) are very focussed on collecting data. We invest an inordinate amount of time and energy in collecting data and we don't think how we could do it differently.... as collecting it as part of a process, as a by-product of a process. We should look at the way we do things and see how we can derive data as a by-product of that process supported by technology” (I11:SS).

In the absence of EHRs the interviewees considered that, where possible, electronic data should be reused to improve efficiency. The preliminary quantitative findings, as discussed in Section 5.2.4, had revealed the possibility that 90% of the required data for the SACT reimbursement scheme were already collected in one software system used in the majority of the hospitals. The interviewees’ views on potential ways to improve this data collection, are detailed in Table 6.2 below, and included options such as reporting the existing electronic data and subsequently submitting any manual data, amalgamation of existing electronic data and any manual data prior to electronic transmission, or implementing a new system

Table 6.2: Interviewees' views on the reuse of existing electronic data

Option 1	Existing electronic data should be electronically extracted and transmitted, and the missing data added manually at a later stage.
Option 2	Amalgamate the missing data with the existing electronic data in an electronic format before transmission, particularly where the missing data might also be available electronically on an existing hospital system, for example integrating the pharmacy system and the HIPE system for the provision of the diagnosis field. *"System integration/linking is key to the provision of data" (I16:SS).* There was an acknowledgement that without 100% of the required data being available electronically there would still be a requirement to go to a paper chart for the remaining information and it would be worth making the investment to get the other 10% electronically.
Option 3	The third option suggested was to investigate new systems, such as EHRs, that would collect all the required data electronically in the process of direct patient care.

Data quality processes are required in the collection of the original source data as well as its reporting for secondary use, as described in Section 2.8.1.2 of the literature review, where data quality is considered a guiding principle in the collection of health data (HIQA, 2013c). The interviewees felt that the data quality assurance processes were undervalued, with little understanding of the requirement to ensure data accuracy and validity to prevent erroneous or incomplete data reporting, particularly where manual entry of data was required, as in the selected case. This was a concern, as the incorrect or incomplete data collected could be used for national reports, policy development as well as funding, leading to biased outcomes and incorrect national funding and service development.

Box 6.3: Interview quotations on data quality

"Low data returns result in a collation of nonsense" (I3:SS)

"Good data is when almost as much as possible is available to study and analyse" (I18:PS).

"The Irish health system does not recognise the profession of people who are responsible for data quality and so on, as a result ensuring the quality of the data is everybody's responsibility and thus nobody's, and there is no clear accountability and responsibility for the data. This is especially pertinent for secondary use, as it is dependent on the data collection and quality implemented for primary use" (I9:SS).

One of the fundamental requirements for data quality is a complete data set (HIQA, 2013c). There was a perception amongst the interviewees that mandatory reporting could resolve some of the data reporting issues, but only where there were penalties for non-compliance. The reporting of data for the SACT reimbursement scheme is not mandatory but does have a financial incentive. Others felt that user-friendly interfaces and incentives were preferable to mandatory reporting.

Box 6.4: Interview quotations on the subject of mandatory reporting

"One positive aspect of mandatory legislation would be that the data would be easier to collect i.e. greater cooperation from the organisations involved. The downside of that is, that where data collection is mandated with legislation, there is much stricter controls on the data collected and its use" (I7:SS).

"Reporting of cancer in the UK is mandatory; having said that, a mandatory requirement may not necessarily guarantee full and accurate reporting, although it is seen by some as "the holy grail". Implementation is very important as, if there are no sanctions for non-compliance, then whether it is mandatory or optional is irrelevant. A better solution would be a robust implementation of a user-friendly data collection process" (I3:SS).

6.4 Challenge - Variations in Measurement Systems

Variations in the disparate datasets being collected lead to requirements for data validation and continuous updating (AHRQ, 2006, HIQA, 2013c) and barriers to the efficient sharing and collection of data.

The comparison of the SACT reimbursement scheme dataset to that of the NCRI and HIPE datasets showed differences in the taxonomies used for the "County" field and also in other areas such as drugs, as discussed in Section 2.8.4.1 and detailed in Appendix 5. Variations in the collection of data in local systems were perceived as more significant by the interviewees, with examples given of variations in fields such as demographics and referral dates with some hospitals collecting "date attended" as first contact of a patient with the hospital and others as the date of receipt of referral letter. Another example given of variations in systems was the challenge presented when attempting to use data for purposes other than for which they were originally collected, resulting in more complex collection and analysis. The example presented was the patient identifiers required by the SACT reimbursement system, being those of primary care such as PPSN, DPS number, or GMS number, rather than those of secondary care, being the hospital identifier. This was further supported by the findings of the quantitative component of this research, as discussed in Section 5.6.

Box 6.5: Interview quotations on variations in measurement systems

> *"One reason for inefficiency is that the same data is required by many reporting systems, mainly the same, but some have variations. This results in duplicate data collection and reporting, whether manual or electronic" (I17:SS).*
>
> *"The main driver for the system (PCRS) is primary care financial management, not outcome orientated. This is a barrier to the implementation of the SACT system" (I4:PS).*

6.5 Challenge - Organisational and Cultural Issues

As discussed in the literature review, it is important that the persons involved in data collection understand the requirement and impact of that data collection (McDevitt and Comber, 2009, Wilkinson et al., 2007). The SACT reimbursement scheme, as described in Chapter 3, was introduced to support the implementation of the government's "money follows the patient" strategy and to centralise the SACT budget, with the additional benefit of providing data to support service planning and clinical audit.

The results of the semi-structured interviews showed that the majority of interviewees understood the requirements for data collection, both in the selected case and the wider domain of public health. They were also aware of national agencies currently collecting data on cancer, including hospitals, the Economic and Social Research Institute, through the HIPE system, the NCRI, the NCCP, in addition to HIQA audits.

A wider cultural issue, namely a lack of information tradition within the HSE, also emerged as a finding of the interview data analysis. This was also considered to be a wider societal issue where anecdote is often believed at the expense of the facts.

Box 6.6: Interview quotations on the requirement for data

> *"Decisions are data driven, or should be data driven (O'Reilly, 2013)".*
>
> *"Fundamentally, to know and be able to analyse and understand what is going on needs good data. The more data you capture in the process then the more capability you have to turn that into useful intelligence in how to improve the process" (I18:PS).*
>
> *"Data is needed to facilitate the commercial management of any large organisation" (I13:SS).*
>
> *"Ireland, not dissimilar from other countries, is not a data-driven society: the anecdote is strong and wins out compared to the population argument"(I3:SS).*

Section 2.8.3 of the literature review revealed that successful implementation of systems, such as data reporting, must be led by senior responsible project owners at local and national level, involving project management principles (Wanna, 2007), and requires openness to change in addition to robust change management programmes, which include review of work practices (Kellermann and Jones, 2013). The findings of the interviews showed that the interviewees recognised that these requirements were important and acknowledged that other concerns, such as lack of ICT infrastructure or staff, would frequently be cited as a barrier to implementation of projects. The quantitative component of this research, as presented in Section 5.2.6, showed that 8 of the 17 respondents had not yet appointed a data manager to the SACT reimbursement scheme data collection. This indicates a deficit in local change management programmes and is at odds with the good governance practice of named persons being responsible for data collection (HIQA, 2013c, AHRQ, 2010). While the requirement for data was acknowledged, it was perceived by the interviewees that the responsibility of the collection and reporting is diluted without the support of dedicated data managers.

Fear was proposed as a challenge by more than one interviewee. One fear described was that of misrepresentation of data leading to reputational damage, where data made available is not attributed with its full meaning, or that meaning is misunderstood, giving a selective picture. The fear of being exposed as inefficient or having poorer outcomes through benchmarking was also raised, which is a potential outcome of clinical audit (O'Reilly, 2013).

Cooperation of the primary stakeholders is essential for project success and to ensure that the primary stakeholders' needs are met, as discussed in Section 2.9 (OGC, 2009, PRINCE2 learning provider ILX Group, 2013). Lack of this cooperation between the NCCP, the PCRS and the hospitals, the primary stakeholders of the SACT reimbursement scheme, was presented as a potential challenge by some interviewees.

Two-way communication between local and national policy leaders is necessary to ensure that challenges to programme implementation are recognised and addressed early in the process (James, 2003). This communication is also necessary to ensure that local benefits are derived from participation in national data collection processes, in addition to maintaining staff motivation for data reporting. It also emphasises the requirement for the data collection, maximises its value through reuse, and secures stakeholder buy-in, as discussed in Sections 2.7 and 2.8.3 (Wilkinson et al., 2007, HIQA, 2013c, Wanna, 2007). It could also address the feeling amongst some of the interviewees that national data collection was fundamentally *"meddling by bureaucrats" (I2:PS)* with no local benefit gained through the provision of these data.

Box 6.7: Interview quotations on change management and incentives

“Human barriers are bigger than ICT (barriers)"(I3:SS).

"Technical barriers can always be overcome if you have the will but when the will is lacking and there are no incentives then the technical barriers are often presented as being the problem when in reality they are not" (I7:SS).

“Everyone thinks it is a very good idea (to collect the data) but no one wants to do it" (I3:SS).

Staff support in the areas of data quality, education and training were also considered important, as discussed in Sections 2.8.1 and 2.8.3. Continued success of implemented programmes requires training, education, and staff recruitment to ensure that more than one person can adequately perform business-critical tasks. Without this redundancy, and with personnel losses, systems can become inoperable, underutilised or abandoned

Box 6.8: Interview quotations on education and training of staff

"There are people who initiate systems, but then there has to be the people who further develop and expand systems, not always the same person;, you need to keep the process of the system alive" (I8:SS).

"High quality data-gathering requires data definitions, minimum data sets and data dictionaries. The challenge is to educate the data managers and the data processors to apply the definitions accurately" (O'Reilly, 2013).

The interviewees also suggested that software systems in use may not be exploited to their full potential due to lack of technical skills of those utilising the system, as discussed in Section 2.8.3, nor implemented in full due to cultural issues or a reluctance to change business processes (AHRQ, 2006). This was also revealed in the quantitative component of this research, as discussed in Section 5.2.5, where some participants had difficulty in extracting data from systems where others did not, and also in the literature, as discussed in Section 2.8.3. Furthermore, the SACT preparation system common to 22 of the 26 hospitals, was revealed as having the potential to record all the data required for the SACT reimbursement scheme, if fully implemented. A work process change could improve efficiency by capturing the required data and supporting its output for secondary use, without the requirement for manual chart abstraction or re-keying of data.

Box 6.9: Interview quotations on challenges to full utilisation of software systems

“Systems are not always utilised to full capacity as users may be skilled to a certain point but not beyond that. As a result, organisations may not be realising the full benefits of the implemented systems. Users often are very competent in operating the system but not in extracting data from the system; for example outpatient booking systems - extremely good for doing what they do, but if a report was required to show average waiting time, there is no one in the hospitals skilled enough to produce that” (I7:SS).

"There needs to be a strategic vision and joined-up thinking. Also there needs to be accountability when health ICT projects do not achieve what they were meant to or could have achieved" (I13:SS).

The findings of Section 2.8.3 of the literature review revealed that accountable leadership at national level, with an understanding of the coalface operation, was essential to successful software system implementation. The quantitative component of this research indicated a deficit in the implementation of software systems that would support the prescribing and administration of SACT, as discussed in Section 5.2.3. The interviewees attributed the lack of progress in implementing such ICT systems in part to the DoH’s restriction on the procurement and implementation of local solutions and the preference for the implementation of national solutions (HSE, 2012a), as discussed in Section 2.8.5, and also to a lack of national leadership to promote the requirement for these systems.

An alternative approach to national ICT system procurement was suggested by the interviewees, where frameworks would dictate the standards required to ensure system interoperability and overcome the perceived lack of system interoperability which precludes communication and data reuse within and between institutions, as discussed in Section 2.8.4.2. This would support the interoperability required for data collection, as discussed in Section 2.6.

Box 6.10: Interview quotations on procurement approaches

“New systems require a project construction that facilitates agreement on requirements; then go to the market place for a framework to meet the requirements of the standards and thus be interoperable" (I11:SS).

There was a general feeling amongst the interviewees, that ICT could support clinical processes and provide the required data efficiently as a by-product of those processes, as discussed in Section 2.6, and that the real barriers were financial and cultural. In addition, a number of the interviewees proposed that fear of technology ageing and decision deferral in the expectation of an ideal system were amongst the motivations delaying progress on ICT implementations.

Box 6.11: Interview quotations on challenges to ICT implementation

“Technology can be sorted if there is money for resources" (I17:SS).

“We are twenty years behind the supermarkets” and “a lot of this (data collection) could be done by information technology” (I12:PS).

“Because we are so behind in understanding ICT supporting processes we invest in data collection, and try and manipulate it and deal with it; this is quite a costly exercise that could be done differently"(I11:SS).

"Like everything, if you are fearful of any development, well then, there is only one way you are going, and that's down" (I8:SS).

"Technology is a challenge and the cost of that technology is part of this. The fact that technology keeps getting better and better, faster and faster makes people tend to hold off until the perfect solution comes along….. which may never come along"(O'Reilly, 2013)

6.6 Challenge - Technological Barriers for Electronic Health Records

Data standards and interoperability are two components essential to data collection and aggregation (Karp et al., 2008, empirica GmbH, 2008, Marcheschi et al., 2005, Stead et al., 2005), as discussed in Section 2.8.4.2 of the literature review. In Ireland, this will require the implementation of the essential building blocks of a national health infrastructure network, such as unique identifiers, as discussed in Section 6.8, and health information standards such as a national dataset for demographics, minimum datasets for diseases, as well as a national drug database, as discussed in Section 2.8.4.2.

Many of the primary and secondary stakeholders were unaware of standards that were already in use or could potentially be used in the area of data collection for secondary use. Some of the secondary stakeholders had previous experience of the use of data standards, minimum data sets and the requirement for data consistency, and were aware of bodies, such as HIPE, the Economic and Social Research Institute (ESRI), the NCRI and the National Treatment Purchase Fund (NTPF) collaborating on the coding and field definitions used in their minimum data sets (I17:SS, II3:SS, I7:SS). This is in line with HIQA's recommendations on data collection, as discussed in Section 2.8.4.2, which emphasises the reuse of existing standards and data definitions (HIQA, 2013c). That notwithstanding, the researcher, as discussed in Section 5.6, found variations in some of the data fields through comparison of the datasets of the NCRI, HIPE and the SACT reimbursement scheme, with one field "county" used as an example.

Interviewees emphasised the immediate requirement to move ahead with standards and agreed on the importance of minimum data sets to ensure interoperability in the future and prevent growing silos of legacy data. HIQA have proposed many of the building blocks required to be in place, as discussed in Sections 2.8.4.1 and 2.8.4.2; but legislation is required to mandate their use.

Box 6.12: Interview quotations on standards and interoperability

> *"If we are going to collect data, at least we should collect it in a structured format. If we don't collect it in a structured format it will be difficult to get any real benefit from it…….All the themes raised are issues and challenge,s but something can be done about them as can be done with regard to variations in measurement systems" (I11:SS).*
>
> *"The lack of system interoperability and communication is a nightmare which is preventing data reuse - within and between institutions" (I4:PS).*

Some interviewees proposed that a sense of realism was required as to what can be achieved within current technical structures, in addition to a pragmatic need for interim solutions while continuing the ongoing search for the ideal system, so as not to completely block all development.

Box 6.13: Interview quotations on technology strategies

> *"Trying to do technology for business in the perfect proper way is pretty much a recipe for not doing it. The other extreme is doing things on such a piecemeal basis without a strategy that all of your efforts are wasted because it is so piecemeal and disjointed. In the middle there is, maybe something like a business strategy being supported by technology, with a focus on delivery, but the strategy is there to ensure that you are not doing things totally piecemeal or totally haphazardly, but at the same time you are delivering and getting things done" (I18:PS).*

6.7 Challenge - Economic Pressures and Competing Priorities

The requirements for data are growing, including the needs of the NCCP, as discussed in Section 2.4. It was accepted by the interviewees, and acknowledged in Section 2.5 of the literature review, that health data are mainly generated at the point of patient care, requiring these institutions to report it for secondary use (Safran et al., 2007, Cimino, 2011). Without the provision of dedicated personnel resources (O'Reilly, 2013) the hospitals involved in the provision of cancer services must organise their workforces to ensure that these data are collected, quality-assured and reported within current personnel complements, absorbing the additional personnel resource requirements that have been quantitatively demonstrated, as discussed in Section 5.6. This leads to competition and requirement for prioritisation with regard to the allocation of finance and staffing (Burke, 2010). There was recognition that competing local and national priorities can result in precedence being given to national rather than local requirements; one example cited a reduction in data collected for cancer research locally, as the time allocated was reassigned to the national reporting priority. Interviewees also felt that there was competition between local clinical and administrative data requirements, leading to local prioritisation of reporting due to the constraints mentioned previously, as well as the need to deal with direct patient care activities as a priority. Ensuring that data is available to be collected and reusing data in so far as possible will minimise this resource requirement (HIQA, 2013c).

Box 6.14: Interview quotations on competing priorities

"Reporting is always going to be bottom of the list of things to do, as it's not seen as the core of anyone's work and not direct patient care" (I2:PS).

"There is a "couldn't be bothered" attitude towards reporting where it is not perceived as a priority" (I3:SS).

"Tracking and quality of service has now overtaken all other considerations" I16:SS.

Data collection is a costly exercise, as discussed in Sections 2.8.1, 2.8.5 and 5.6, the value of which could be maximised through data reuse (Cimino, 2011, HIQA, 2013c). The lack of financial resources to support overtime payments, ICT development and interfacing was mentioned by many of the interviewees, and financial incentives such as the government's policy of "money follows the patient" are considered key to the success of data reporting. The financial reimbursement following from the data collection of the SACT reimbursement scheme was seen by the interviewees as an incentive as it improves the hospitals' financial positions, while also facilitating equitable funding of drugs and removing post-code treatment lotteries, as discussed in Section 2.3.

Box 6.15: Interview quotations on cost of data collection

"It does not matter what area you are talking about, data is expensive to collect and you really want to make as much use of it as you can for secondary purposes as well as primary purposes" (I9:SS) .

The interviewees suggested additional fields which, if collected, could make the SACT reimbursement scheme dataset more meaningful. These included the treating consultant, treatment intent, patient catchment area, adjuvant or neoadjuvant treatment, tumour staging designations, disease presentation such as relapsed or refractory disease, in addition to a local hospital identifier, such as MRN, to facilitate local identification of patients. These would be similar to the dataset collected in countries operating a national cancer information system (NHS CIU, 2011), as discussed in Section 2.4.1. To counter this, there was a pragmatic awareness of the need for a clear vision of data requirements, which considers the ease or difficulty with which the required data can be collected, and whether the new data would be useful, this view was similar to the findings of the literature review and emerging recommendations on evaluating the requirement for data collection (HIQA, 2013c).

Box 6.16: Interview quotations on minimum data set requirements

"It behoves the organisation that requests data that a) they really need it, b) it has a purpose and c) they use it for that purpose"(O'Reilly, 2013).

“What is collected has to make sense, you need to collect information that is useful and used” (I8:SS).

“We need to analyse data already collected for its strengths, weaknesses and usefulness. If what we collect is not used then collecting more is not the answer" (I3:SS).

A number of interviewees went beyond the challenges of the collection of the data, to the need for business intelligence, combining technical and clinical skills, and to model the data in a way that enables clinical users to extract it meaningfully, as discussed in Section 2.7. The interviewees perceived that there was a lack of these skills at national level, resulting in the collection of data that is not utilised or does not have its potential exploited both nationally and locally, as also discussed in Section 2.7.

Box 6.17: Interview quotations on business intelligence

"People don't know how to use data, but continue to collect it" (I12:PS).

"Where data is gathered there is a fundamental issue with organisation and making the best use of the information". "Data is a technical thing but clinical data needs clinical interpretation, needs a data source, needs people who can work the data source and clinical people inputting on what is needed" (I18:PS).

“Any data collection is obviously going to be beneficial to somebody; depends on what you do with. It’s like everything else, it’s like a hammer - a hammer is a tool, and it’s a fabulous tool - but it is no good unless you pick it up and bang a nail with it"(I6:SS).

6.8 Data Ownership, Data Protection and Identifiers

The introduction of the Health Information Bill has been on the DoH agenda since close of consultation in 2008 (DOH, 2013b). As discussed in Sections 2.8.4.1 and 2.8.6.1 of the literature review, the introduction of this bill is anticipated to facilitate the implementation of standards based unique health identifiers for individuals, healthcare practitioners and organisations and to balance the rights of an individual to have their personal health information protected with the requirement to have an improved health service (Bloomrosen and Detmer, 2008). In the absence of legislation there are difficulties, particularly with regard to collection of population data for secondary use, which is severely hindered by the requirement to obtain informed consent from each patient involved (Data Protection Commissioner, 2013a, Data Protection Commissioner, 2013b).

The findings of the thematic analysis of the interviews reflected those of the literature review, where the subjects of data protection and patient consent were raised with regard to the privacy of individuals. The interview findings indicated that interpretation of data protection legislation, in the domain of Irish healthcare, is highly individualised and conservative. Many interviewees referred to the recent effort required to prevent the HSE from destroying the genetic data of 1.5 million children, following a single complaint, on the advice of the DPC (Crowe, 2013, Corbet, 2013). The balancing of the benefit to society, versus the risk to an individual, if the samples were preserved for medical research, did not seem to be considered. This is analogous to the tension between the rights of an individual and the requirements of public health, as presented in the case for data stewardship in Section 2.4.

A barrier cited by the interviewees as atypical, in contrast to other secondary use hospital data collections, was the requirement to report SACT reimbursement scheme data on a named patient basis. This differs from the other national secondary care data collections, such as HIPE and the NTPF, which are anonymised (HIQA, 2013a), yet is similar to the NCRI data collection requirements. The implementation of the government’s “money follows the patient” funding model will increase the requirement for patient-identifiable data, as discussed in Chapter 3.

The findings of the qualitative data indicated that legislation was not seen as obviating the requirement to inform patients as to the use and potential future use of their health data; rather, it was recognised that there is a requirement to educate patients, and the wider public, to the requirement for data sharing for the greater good and societal benefit, in addition to providing reassurance as to how their data will be protected in line with the recommendations on data stewardship, as presented in Section 2.8.6.1 of the literature review. That notwithstanding, the majority of interviewees believed that with the correct safeguards in place no harm would be done to patients through the reuse of their data.

Box 6.18: Interview quotations on data protection

"Ireland operates with a highly conservative interpretation of data protection legislation for healthcare, it’s fine to take a load of information on someone shopping in Tesco but you cannot share information between out-patients and GPs" (I3:SS).

"Secondary use would not be detrimental to patients if used confidentially. People (the public) are too sensitive about the collection of these (health) data."(I12:PS).

The Irish healthcare sector currently operates in the absence of standardised and mandated healthcare identifiers, as discussed in Section 2.8.4.1. The lack of legislation to support the sharing and linkage of patient information, as well as the absence of unique identifiers for patients, hospitals and providers, was frequently mentioned as a challenge by the interviewees, which is pertinent as discussed in Sections 2.8.6.1 and 2.8.4.1.

In addition, the barriers to the use of the PPSN as an identifier was raised by a number of interviewees, in line with the finding of the literature review, as discussed in section 2.8.4.1, where the DPC does not advocate the use of the PPSN for health identification (Davis, 2009).

Box 6.19: Interview quotations on unique identifiers

"The unique identifier would be handy, but unless there is legislation introduced to allow it to be used to link data, it is useless. The unique identifier will only work if underpinned by legislation to link datasets; the Health Information Bill will need to be introduced in its totality to be useful" (I7:SS).

"The lack of a universal patient identifier presents a significant challenge to the linkage of data between any systems outside of the treating hospital" (O'Reilly, 2013).

"The PPSN is not to be used without consent, yet is key to the PCRS dataset for patient identification" (I17:SS).

Data ownership issues presented strongly in the interview findings, especially in relation to existing clinical databases or systems specifically designed for local use. This was acknowledged to be exacerbated by the lack of a data-sharing culture and the perception that "*everyone who needs to know about my data knows" (I3:SS)* without sometimes considering the broader picture, including the benefits that data could bring if shared at a national level. Reluctance to share data due to data ownership questions and fear of the misinterpretation of shared data emerged as additional challenges during the course of the research, highlighting again the requirement for legislation in addition to a data stewardship model, as discussed in Section 2.8.6.1.

Box 6.20: Interview quotations on data ownership

"It is my data" and "I built and designed this database" (I3:SS referring to another person).

6.9 Conclusion

This chapter presents the analysis of the narrative data of the semistructured interviews and the overall findings of the research, integrating the findings of the interviews, the literature review and the quantitative component of this research. The findings are presented in themes as established by the AHRQ, and answer the questions posed in Chapter 1 as to the contemporary Irish challenges and facilitators to the collection of health data for secondary use. The requirement for additional personnel had been described in the quantitative component of this research, in addition to the finding of the literature review in Chapter 2, and was further substantiated by the accounts of the interviewees.

The key conclusions of this research and associated recommendations are presented in Chapter 7.

Chapter 7. Conclusion

The objective of this research was to describe the context for collection of health data for secondary use in Ireland, to identify actual and perceived barriers to the collection of these data, and to point out opportunities that would facilitate that collection, in addition to quantitatively ascertaining the personnel resources required to report the data of the selected case. The determination of these challenges provides the evidence necessary to inform prioritisation of the steps and projects essential to improving this phenomenon.

The results of the research and associated findings have been presented and discussed in the preceding chapters. This chapter contains a short account of these findings, the researcher's conclusions, recommendations for future research and the study limitations.

7.1 Research Summary

Initial research and the literature review indicated that the secondary use of health data is topical and recognised as essential for service improvement, governance, research, funding, and government policy. The existence of challenges to the secondary use of data is well recognised internationally. These challenges have been subcategorised into six discrete areas by the AHRQ, and the AMIA have proposed a framework and data stewardship model to help overcome some of these challenges.

A case study was undertaken following from the literature review, into a single case, the SACT reimbursement scheme, a recently implemented requirement for additional data reporting to support secondary uses of health data in the areas of reimbursement and clinical audit. The personnel resource aspect of the selected case was further researched, as an embedded subunit of the larger case study. The purpose of the case study was to reveal the actuality of any challenges and in doing so allow an opportunity for reflection and the potential for a new direction in future developments.

7.2 Implications of the Research Findings

The findings of the research have both practical and theoretical importance as they can serve as the basis for future research on the subject, and can also aid in improving the practices related to the research problem, while enabling a number of recommendations as described in Section 7.4 below.

7.3 Limitations of the Research Findings

Many of the issues, described in this research, were communicated to me by interviewees, with supporting information from the literature review, in addition to the quantitative component of this research. In the circumstances of the selected case, the validity of my findings are not in question, but the unitary nature of the case makes me cautious about generalising more widely. In support of the generalisability of the study, however, other Irish data collections, in comparable circumstances, could be expected to encounter similar challenges. The main limitation of my study is that I may have missed important factors because they were not present in the selected case or in the participating hospitals. Therefore, I cannot conclude that the challenges highlighted in this dissertation are the only or most important ones.

It also has to be stated that the research was conducted in an Irish context. The challenges encountered in other jurisdictions may vary from those described, due to sociocultural and technological differences.

The reader may also wish to consider the potential personal biases of the researcher, as detailed in Section 4.3.1.1, including the author's self-declared worldview of pragmatic constructivism, may well have influenced the direction of this research. In addition to these factors there is the author's self-interest in seeking to achieve an MSc in Health Informatics through completion of this dissertation.

7.4 Research Conclusions and Recommendations

Secondary use of health data is a complex phenomenon which faces numerous challenges, including variation in measurement systems, organisational, cultural, technical and economic issues, in addition to competing priorities.

Evidence from this research, conducted in the domain of secondary care, indicates that data collection for secondary use is undertaken as a process separate from the collection of data in the course of the clinical care of the patient, resulting in a requirement for increased resources, including personnel, which may not be readily available in the current economic climate, and must compete with local priorities, including direct patient care, for these limited resources. Data collection programmes require implementation in a manner that is cognisant of the complexity of the data collection in addition to the local economic pressures and staffing resources. These programmes should provide incentives in the form of resource support, as well as meaningful data reports to support local requirements. Collection of health data must be integrated into the documentation of the clinical care pathway of the patient, with minimal disruption, if it is to remain sustainable while accommodating the increasing requirements of a well-managed health service. Data collections require regular review to ensure that the data collected is still meeting the reporting need.

This research has revealed a deficit in Ireland’s ICT infrastructure, preventing the collection of data in an efficient manner, such as the “collect once and use often” paradigm, resulting in a requirement for increased resources. Technology, in the form of electronic health records, has been shown to improve the quality and efficiency of data capture for both primary and secondary uses of these data. However, implementation of these systems requires a new procurement paradigm, with national and local accountable leadership, that specifics minimum interoperability and data reporting requirements, in addition to clinical process requirements, to support a system framework rather than a unitary national system approach.

Data collections should conform to Irish legislative requirements as well as the Health Information Quality Authority’s healthcare information and interoperability standards, and best practice guidelines. This will ensure standardisation of data collection and allow data collected to be shared and adapted, both at local and national level, going forward. However, this standardisation is dependent on the implementation of unique identifiers, standard national datasets for demographics and disease states, as well as a national drug database; which in turn is dependent on the enactment of the Health Information Bill to provide the legal framework for identifiers, data sharing, patient consent and data stewardship.

Improvement in technology is not the only factor that will facilitate an improvement in data collection. The sociocultural challenges associated with change management and programme implementation are equally important. A comprehensive and well-managed change management process in addition to accountable leadership, both at local and national level is necessary to drive review of work practices and ensure that software systems are fully implemented in line with national objectives.

A national approach to healthcare data collection and aggregation, such as a national cancer information system, would eliminate the requirement for multiple data collections within the domain of cancer control. The coordination of the disparate data collections to support the development of a national cancer information system could be a pragmatic approach to a longer-term objective.

7.5 Recommendations for Further Research

The results of this research indicate that there are many challenges encountered in Ireland to the secondary use of health data. Some of these could be overcome by the introduction of a framework, or tool kit, for the implementation of national data collections. Such a framework would include assessment of the data required and estimation of the resource demands generated by new and extended data collections. Such toolkits have been developed in other countries; however they may need adaption for successful operation in an Irish context.

References

1992. Evidence-Based Medicine: A New Approach to Teaching the Practice of Medicine. *JAMA,* 268**,** 2420-2425.

AGRAWAL, A. 2002. Return on investment analysis for a computer-based patient record in the outpatient clinic setting. *J Assoc Acad Minor Phys,* 13**,** 61-5.

AHRQ 2006. Collecting and Reporting Data for Performance Measurement: Moving Toward Alignment. *AHRQ Conference on Health Care Data Collection and Reporting.*

AHRQ 2010. Data Collection and Quality Assurance. *In:* GLIKLICH, R. E. & DREYER, N. A. (eds.) *Registries for Evaluating Patient Outcomes: A User's Guide.* 2nd ed.: Rockville (MD).

ALTARUM INSTITUTE 2011. Overcoming Challenges to Health IT Adoption in Small, Rural Hospitals. Office of the National Coordinator for Health Information Technology.

ANCKER, J. S., SHIH, S., SINGH, M. P., SNYDER, A., EDWARDS, A. & KAUSHAL, R. 2011. Root causes underlying challenges to secondary use of data. *AMIA Annu Symp Proc,* 2011**,** 57-62.

ASHP 2008. The ASHP Discussion Guide on USP Chapter <797> for Compounding Sterile Preparations. American Society of Health-System Pharmacists.

AYATOLLAHI, H., BATH, P. & GOODACRE, S. 2009. Paper-based versus computer-based records in the emergency department: Staff preferences, expectations, and concerns. *Health Informatics Journal,* 15**,** 199-211.

BALANDA, K., MCMASTER, C., PURSELL, L., KEARNS, N., GRAHAM, A., BARRON, S., FAHY, L. & O’DONOVAN, D. 2008. All-Ireland Health and Wellbeing Data Set (AIHWDS) INIsPHO. Dublin.

BARTON, C., KALLEM, C., VAN DYKE, P., MON, D. & RICHESSON, R. 2011. Demonstrating "collect once, use many"-assimilating public health secondary data use requirements into an existing Domain Analysis Model. *AMIA Annu Symp Proc,* 2011**,** 98-107.

BAXTER, P. & JACK, S. 2008. Qualitative Case Study Methodology: Study Design and Implementation for Novice Researchers. *The Qualitative Report,* 13**,** 544-559.

BC POPULATION HEALTH AND WELLNESS 2005. Public Health Renewal in British Columbia: An Overview of Core Functions In Public Health. Ministry of Health Services, British Columbia.

BLOOMROSEN, M. & DETMER, D. 2008. Advancing the framework: use of health data - a report of a working conference of the American Medical Informatics Association. *J Am Med Inform Assoc,* 15**,** 715-22.

BRADLEY, E. H., HERRIN, J., ELBEL, B., MCNAMARA, R. L., MAGID, D. J., NALLAMOTHU, B. K., WANG, Y., NORMAND, S., SPERTUS, J. A. & KRUMHOLZ, H. M. 2006. Hospital quality for acute myocardial infarction: Correlation among process measures and relationship with short-term mortality. *JAMA: The Journal of the American Medical Association,* 296**,** 72-78.

BROWMAN, G. P., BROUWERS, M., FEVERS, B. & SAWKA, C. 2010. Population-based cancer control and the role of guidelines - towards a systems approach. *Cancer Control.* 1 ed.: Oxford University Press.

BURKE, S. 2010. Boom to bust its impact on Irish health policy and health services. *Irish Journal of Public Policy,* 2010.

CALLEN, J. L., BRAITHWAITE, J. & WESTBROOK, J. I. 2008. Contextual implementation model: a framework for assisting clinical information system implementations. *J Am Med Inform Assoc,* 15**,** 255-62.

CARON, L. 2010. Organizational structures for cancer control. *Cancer Control.* 1 ed.: Oxford University Press.

CARRINGTON, C., STONE, L., KOCZWARA, B., SEARLE, C., SIDEROV, J., STEVENSON, B., MICHAEL, M., HYDE, S., BOOTH, A. & RUSHTON, S. 2010. The Clinical Oncological Society of Australia (COSA) guidelines for the safe prescribing, dispensing and administration of cancer chemotherapy. *Asia-Pacific Journal of Clinical Oncology,* 6**,** 220-237.

CIMINO, J. J. 2007. Collect Once, Use Many: Enabling the Reuse of Clinical Data through Controlled Terminologies. *Journal of AHIMA,* 78**,** 24-29.

CIMINO, J. J. 2011. High-quality, standard, controlled healthcare terminologies come of age. *Methods Inf Med,* 50**,** 101-4.

COLEMAN, M. P., FORMAN, D., BRYANT, H., BUTLER, J., RACHET, B., MARINGE, C., NUR, U., TRACEY, E., COORY, M., HATCHER, J., MCGAHAN, C. E., TURNER, D., MARRETT, L., GJERSTORFF, M. L., JOHANNESEN, T. B., ADOLFSSON, J., LAMBE, M., LAWRENCE, G., MEECHAN, D., MORRIS, E. J., MIDDLETON, R., STEWARD, J. & RICHARDS, M. A. 2011. Cancer survival in Australia, Canada, Denmark, Norway, Sweden, and the UK, 1995-2007 (the International Cancer Benchmarking Partnership): an analysis of population-based cancer registry data. *The Lancet,* 377**,** 127-138.

COMBER, H. Secondary use of data - striking a balance Promoting Health Research and Protecting Patient Rights, 2006a.

COMBER, H. 2006b. Trends in Irish cancer incidence 1994-2002, with projections to 2020. National Cancer Registry (2006).

CORBET, R. 2013. Expert Comment. *Data Protection Ireland,* 6.

CRESSWELL, K., MORRISON, Z., CROWE, S., ROBERTSON, A. & SHEIKH, A. 2011. Anything but engaged: user involvement in the context of a national electronic health record implementation. *Inform Prim Care,* 19**,** 191-206.

CRESWELL, J. W. 2009. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, SAGE Publications.

CROWE, J. 2013. Plan to destroy heelprick cards is a serious mistake. *The Irish Times*, 19 March 2013.

CUSACK, C. M., HRIPCSAK, G., BLOOMROSEN, M., ROSENBLOOM, S. T., WEAVER, C. A., WRIGHT, A., VAWDREY, D. K., WALKER, J. & MAMYKINA, L. 2012. The future state of clinical data capture and documentation: a report from AMIA's 2011 Policy Meeting. *JAMIA*.

DATA PROTECTION COMMISSIONER. 2013a. *A Guide for Data Controllers* [Online]. Available: http://www.dataprotection.ie/ViewDoc.asp?fn=/documents/guidance/Guide_Data_Contollers.htm&CatID=90&m=y [Accessed 6 June 2013].

DATA PROTECTION COMMISSIONER. 2013b. *Law on Data Protection* [Online]. Available:

http://www.dataprotection.ie/ViewDoc.asp?fn=%2Fdocuments%2Flegal%2FLawOnDP%2Ehtm&CatID=7&m=l [Accessed 8 March 2013].

DATA PROTECTION COMMISSIONER. 2013c. *Public Register* [Online]. Data Protection Commissioner. Available: https://www.dataprotection.ie/ViewDoc.asp?fn=/documents/register/default.asp&CatID=27&m=g [Accessed 7 March 2013].

DAVIS, G. 2009. Submission from Deputy Commissioner, Gary Davis on behalf of the Office of the Data Protection Commissioner regarding the proposed Health Information Bill. Data Protection Commissioner.

DEPARTMENT OF HEALTH 1998. Information for health: an information strategy for the modern NHS 1998-2005. London: NHS Executive.

DESHMUKH, V. G., SOWER, N. B., HUNTER, C. Y. & MITCHELL, J. A. 2011. Integrating historical clinical and financial data for pharmacological research. *BMC Med Res Methodol,* 11**,** 151.

DOH 2004. Health Information: A National Strategy. Department of Health & Children.

DOH 2008. Building a Culture of Patient Safety: Report of the Commission on Patient Safety and Quality Assurance. Department of Health & Children.

DOH 2013a. Money Follows the Patient - Policy Paper on Hospital Financing. Department of Health & Children.

DOH. 2013b. *Proposed Health Information Bill* [Online]. Department of Health & Children. Available: http://www.dohc.ie/issues/hib/ [Accessed 16 Feb 2013].

DUFFY, J. 2009. Circular 2/09: Arrangements for ICT Expenditure in the Civil and Public Service. Department of Finance.

EHR4CR 2011. Executive Summary. Electronic Health Records for Clinical Research.

EL FADLY, A., RANCE, B., LUCAS, N. L., MEAD, C., CHATELLIER, G., LASTIC, P.-Y., JAULENT, M.-C. & DANIEL, C. 2011. Integrating clinical research with the Healthcare Enterprise: From the RE-USE project to the EHR4CR platform. *Journal of Biomedical Informatics,* 44, Supplement 1**,** S94-S102.

ELWOOD, J. M. & SUTCLIFFE, S. B. 2010. Cancer control and the burden of cancer. *Cancer Control.* 1 ed.: Oxford University Press.

EMPIRICA GMBH 2008. ICT standards in the health sector: current situation and prospects. *epractice.eu.*

FINK, A. 1993. *Evaluation Fundamentals: Guiding Health Programs, Research, and Policy,* Sage Production.

FINK, A. 2009. *Conducting Research Literature Reviews: From the Internet to Paper*, SAGE Publications.

FREEMAN, E. 1984. *Strategic Management: A Stakeholder Approach*, Cambridge University Press.

GAO 2010. Higher Education: Institutions' Reported Data Collection Burden Is Higher Than Estimated but Can Be Reduced through Increased Coordination. United States Government Accountability Office: Report to Congressional Committees.

GARRARD, J. 2010. *Health Sciences Literature Review Made Easy: the Matrix Method*.

GELMON, S. B. & DROPPERS, O. 2008. Community Health Centers and EHR Records: Challenges and Opportunities Northwest Health Foundation.

GOMM, R., HAMMERSLEY, M. & FOSTER, P. 2000. *Case Study Method: Key Issues, Key Texts* Sage Publications Ltd.

GROOVER, M. P. 2007. *Work Systems and Methods, Measurement, and Management of Work*, Pearson Education International.

GROSSMAN, J. M., KUSHNER, K. L. & NOVEMBER, E. A. 2008. Creating sustainable local health information exchanges: can barriers to stakeholder participation be overcome? *Res Brief***,** 1-12.

HEALTH INDUSTRIES 2009. Transforming healthcare through secondary use of health data. PriceWaterhouseCoopers.

HERNANDEZ, P., PODCHIYSKA, T., WEBER, S., FERRIS, T. & LOWE, H. 2009. Automated mapping of pharmacy orders from two electronic health record systems to RxNorm within the STRIDE clinical data warehouse. *AMIA Annu Symp Proc*, 2009**,** 244-8.

HIMSS 2010. Definition of a Standard. HiMSS Dictionary, 2nd Edition.

HIPE 2013. HIPE Data Dictionary 2013. 5 ed.: ESRI.

HIQA 2007. National Quality Assurance Standards for Symptomatic Breast Disease Services: Developing Quality Care for Breast Services in Ireland. Health and Quality Authority.

HIQA 2009. Recommendations for a Unique Health Identifier for Individuals in Ireland. Health Information and Quality Authority.

HIQA 2011a. Developing National Health Interoperability Standards for Ireland: A Consultation Document. Health Information and Quality Authority.

HIQA 2011b. Draft Standards for National Health Information Resources. Health Information and Quality Authority.

HIQA 2011c. Recommenations for Unique Health Identifiers for Healthcare Practitioners and Organisations. Health Information and Quality Authority.

HIQA 2012a. EPrescribing and Electronic Transfer of Prescriptions: an International Review.

HIQA 2012b. Guidance on information governance for health and social care services in Ireland. Health Information and Quality Authority.

HIQA 2012c. International Review of Secondary Use of Personal Health Information. Health Information and Quality Authority.

HIQA 2012d. National Standards for Safer Better Healthcare. Health Information and Quality Authority.

HIQA 2012e. What you should know about Data Quality: A guide for health and social care staff. Health Information and Quality Authority.

HIQA. 2013a. *Catalogue of National Health Information Sources* [Online]. Available: http://www.hiqa.ie/resource-centre/professionals/health-information-sources [Accessed 6 June 2013].

HIQA 2013b. Draft National Standard Demographic Dataset and Guidance for use in health and social care settings in Ireland. Health Information and Quality Authority.

HIQA 2013c. Guiding Principles for National Health and Social Care Data Collections. Health Information and Quality Authority.

HSE 2011. METADATA: Based on National Service Plan 2011. Corporate Planning and Performance Unit.

HSE 2012a. Health Service Executive: National Financial Regulation: ICT Funding and Approval - NFR-07. 2 ed.

HSE 2012b. Health Service Executive: National Service Plan 2012. Corporate Planning and Corporate Performance Directorate.

HSE 2012c. Health Service Executive: September 2012 Performance Report: National Service Plan 2012. Health Services Executive.

HSE. 2012d. *Health Services Executive - Our Structure* [Online]. Available: http://www.hse.ie/portal/eng/about/Who/Our_Structure.html [Accessed 12 Nov 2012].

HSE. 2012e. *National Cancer Control Programme* [Online]. Available: http://www.hse.ie/eng/services/Find_a_Service/National_Cancer_Control_Programme/ [Accessed 12 Nov 2012].

HSE. 2013a. *2013 Performance Reports* [Online]. Available: http://www.hse.ie/eng/services/Publications/corporate/performancereports/2013pr.html [Accessed 24 June 2013].

HSE 2013b. HSE National Service Plan 2013.

HSE. 2013c. *Planning and Performance Reporting* [Online]. Available: http://www.hse.ie/eng/about/Who/Board_Members/codeofgovernance/planningandperformancereporting.html [Accessed 24 June 2013].

IAKOVIDIS, I. 1998. Towards personal health record: current situation, obstacles and trends in implementation of electronic healthcare record in Europe. *International Journal of Medical Informatics,* 52**,** 105-115.

IOM 2001a. *Crossing the Quality Chasm: A New Health System for the 21st Century*, The National Academies Press.

IOM 2001b. Envisioning the National Health Care Quality Report. Institute of Medicine.

IOM 2003. *Key Capabilities of an Electronic Health Record System: Letter Report*, The National Academies Press.

IOM 2004. *Patient Safety: Achieving a new standard of care*.

IOM 2007. The Learning Healthcare System. Institute of Medicine.

IOM 2009. Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement. *In:* MEDICINE, I. O. (ed.). Washington, DC: The National Academies Press.

ISMP. Guidelines for SAFE Preparation of Sterile Compounds. ISMP Sterile Preparation Compounding Safety Summit, 2013.

ISO 2005. Health informatics — Electronic health record — Definition, scope and context. : International Organization for Standardization, Geneva, Switzerland.

JAMES, O. 2003. *The Executive Agency Revolution in Whitehall: Public interest versus bureau-shaping perspectives,* Palgrave Macmillan.

JONES, R. B., REEVES, D. & MARTINEZ, C. S. 2012. Overview of electronic data sharing: why, how, and impact. *Curr Oncol Rep,* 14**,** 486-93.

KARP, D. R., CARLIN, S., COOK-DEEGAN, R., FORD, D. E., GELLER, G., GLASS, D. N., GREELY, H., GUTHRIDGE, J., KAHN, J., KASLOW, R., KRAFT, C., MACQUEEN, K., MALIN, B., SCHEUERMAN, R. H. & SUGARMAN, J. 2008. Ethical and practical issues associated with aggregating databases. *PLoS Med,* 5**,** e190.

KELLERMANN, A. L. & JONES, S. S. 2013. What it will take to achieve the as-yet-unfulfilled promises of health information technology. *Health Aff (Millwood),* 32**,** 63-8.

KERNAN, B. 2012. *A Medicines and Drugs Reference Catalogue for e-Prescribing in Ireland.* M.Sc. Health Informatics, Trinity College Dublin.

KUSH, R., ALSCHULER, L., RUGGERI, R., CASSELLS, S., GUPTA, N., BAIN, L., CLAISE, K., SHAH, M. & NAHM, M. 2007. Implementing Single Source: the STARBRITE proof-of-concept study. *J Am Med Inform Assoc,* 14**,** 662-73.

KVALE, S. & BRINKMANN, S. 2009. *InterViews: learning the craft of qualitative research interviewing* Sage Publications Ltd.

LOCKE, L. F., SILVERMAN, S. J. & SPIRDUSO, W. W. 2007. *Proposals that work : a guide for planning dissertations and grant proposals,* Thousand Oaks, Sage Publications.

LONGSTAFF, D. 2005. Contentious crop: harvesting information from electronic health records.

LOWE, H. J., FERRIS, T. A., HERNANDEZ, P. M. & WEBER, S. C. 2009. STRIDE - An integrated standards-based translational research informatics platform. *AMIA Annu Symp Proc,* 2009**,** 391-5.

LYNN, J. 2013. *Benefits of EMR or EHR Over Paper Charts* [Online]. Available: http://www.emrandhipaa.com/benefits-of-emr-or-ehr-over-paper-charts/ [Accessed 6 June 2013].

MAGAGNOTTI, N. & SPINELLI, R. 2012. Good practice guidelines for biomass production system. *WG 2 Operations research and measurement methodologies.* Italy: COST.

MARCHESCHI, P., MAZZARISI, A., DALMIANI, S. & BENASSI, A. 2005. New Standards for Cardiology Report and Data Communication: An Experience with HL7 CDA Release 2 and EbXML. *Computers in Cardiology.*

MCDEVITT, J. & COMBER, H. 2009. A Proposed Core National Cancer Dataset: National Cancer Registry. Cork: National Cancer Registry Ireland.

MCDONALD, C. J. 1997. The barriers to electronic medical record systems and how to overcome them. *JAMIA,* 4**,** 213-21.

MCKINSEY 2011. Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute.

MCLEAN, V. 2006. Electronic Health Records Overview. National Center for Research Resources - MITRE.

MENACHEMI, N. & COLLUM, T. H. 2011. Benefits and drawbacks of electronic health record systems. *Risk Management and Healthcare Policy,* 4**,** 47-55.

MILLER, R. H., WEST, C., BROWN, T. M., SIM, I. & GANCHOFF, C. 2005. The value of electronic health records in solo or small group practices. *Health Aff (Millwood),* 24**,** 1127-37.

NATIONAL CANCER FORUM 2006. A strategy for cancer control in Ireland. Dublin: Department of Health.

NCAG 2009. Chemotherapy Services in England: Ensuring quality and safety. National Chemotherapy Advisory Group.

NCCP 2010. Key Performance Indicators Report for Symptomatic Breast Disease Services, 2010. HSE.

NCCP 2012a. ODMS Specification Document. 1.1 ed.

NCCP 2012b. Oncology Drug Management Web Application.

NCIN. 2013. *About the National Cancer Intelligence Network* [Online]. Available: http://www.ncin.org.uk/about_ncin/default.aspx [Accessed 17 Feb 2013].

NCRI 2013. The NCRI Data Dictionary. Available on request from the NCRI.

NEHTA. 2012. *Healthcare Identifiers* [Online]. National E-health Transition Authority. Available: http://www.nehta.gov.au/connecting-australia/healthcare-identifiers [Accessed 4 Jan 2013].

NHS. 2013. *The NHS dictionary of medicines and devices (dm+d)* [Online]. Available: http://www.dmd.nhs.uk/ [Accessed 6 May 2013].

NHS CIU 2011. SACT Dataset 15 ed.

Author. "Title." Date

O'REILLY, S. & VENKATESH, J. 2010. Managing the cost of new therapies: the challenge of funding new drugs. *Cancer Control.* 1 ed.: Oxford University Press.

OGC 2009. *An Introduction to Prince2^{TM}: Managing and Directing Successful Projects* London, The Stationary Office.

ONC 2010. ONC. HIT Standards Committee Draft Transcript. U.S. Department of Health & Human Services.

OSHEROFF, J., PIFER, E., TEICH, J., SITTIG, D. & JENDERS, R. 2005. *Improving Outcomes with Clinical Decision Support:An Implementer's Guide*, HiMSS.

OTTOSEN, T. 2012. EMRs/EHRs: Benefits and Further Information *National Network of Libraries of Medicine (NN/LM) Southeastern/Atlantic Region (SE/A)*

PEVNICK, J. M., CLAVER, M., DOBALIAN, A., ASCH, S. M., STUTMAN, H. R., TOMINES, A. & FU, P. J. 2012. Provider Stakeholders' Perceived Benefit from a Nascent Health Information Exchange: A Qualitative Analysis. *Journal of Medical Systems*, 36**,** 601-613.

PHAM, H. H., COUGHLAN, J. & O'MALLEY, A. S. 2006. The Impact Of Quality-Reporting Programs On Hospital Operations. *Health Affairs,* 25**,** 1412-1422.

PORTER, S. 2013. *Office of the National Coordinator for Health IT Tackles Pressing EHR Issues* [Online]. American Academy of Family Physicians. Available: http://www.aafp.org/news-now/practice-professional-issues/20130612fmostashari.html [Accessed 28 June 2013].

PRINCE2 LEARNING PROVIDER ILX GROUP. 2013. *How to Identify stakeholders* [Online]. Available: http://blog.prince2.com/2011/09/how-to-identify-stakeholders/ [Accessed 4 April 2013].

PWC 2012. Needles in a haystack: Seeking knowledge with clinical informatics. Health Research Institute.

REA, S., PATHAK, J., SAVOVA, G., ONIKI, T. A., WESTBERG, L., BEEBE, C. E., TAO, C., PARKER, C. G., HAUG, P. J., HUFF, S. M. & CHUTE, C. G. 2012. Building a robust, scalable and standards-driven infrastructure for secondary use of EHR data: the SHARPn project. *J Biomed Inform,* 45**,** 763-71.

REILLY, J. 2011. Infectious Diseases (Amendment) Regulations 2011. *S.I. No. 452/2011.* Office of the Attorney General.

RICHARDS, L. 2008. *Handling Qualitative date: A Practical Guide,* Sage Publications.

RICHARDS, M. 2010. Improving cancer services: the approach taken in England. *Cancer Control.* 1 ed.: Oxford University Press.

SAFRAN, C., BLOOMROSEN, M., HAMMOND, W. E., LABKOFF, S., MARKEL-FOX, S., TANG, P. C. & DETMER, D. E. 2007. Toward a national framework for the secondary use of health data: an American Medical Informatics Association White Paper. *J Am Med Inform Assoc,* 14**,** 1-9.

SHAPIRO, J. S. 2007. Evaluating public health uses of health information exchange. *Journal of Biomedical Informatics,* 40**,** S46-S49.

SHEKELLE, P. G., MORTON, S. C. & KEELER, E. B. 2006. Costs and benefits of health information technology. *Evid Rep Technol Assess.* 2007/07/14 ed.: AHRQ.

SHPA 2005. SPHA Standards of Practice for the Safe Handling of Cytotoxic drugs in Pharmacy Departments. *Journal of Pharmacy Practice and Research,* 35**,** 44-52.

SIDOROV, J. 2006. It Ain't Necessarily So: The Electronic Health Record And The Unlikely Prospect Of Reducing Health Care Costs. *Health Aff (Millwood),* 25**,** 1079-85.

SILOW-CARROLL, S., EDWARDS, J. N. & RODIN, D. 2012. Using Electronic Health Records to Improve Quality and Efficiency: The Experiences of Leading Hospitals. The Commonwealth Fund.

STEAD, W. W., KELLY, B. J. & KOLODNER, R. M. 2005. Achievable steps toward building a National Health Information infrastructure in the United States. *J Am Med Inform Assoc,* 12**,** 113-20.

STROETMANN, K. A., JONES, T., DOBREV, A. & STROETMANN, V. N. 2006. eHealth is Worth it. European Commission Information Society and Media DG.

TURNER, M. 2009. HEALTHSTAT: MEASURING THE PERFORMANCE OF THE IRISH PUBLIC HEALTH SERVICE. *Journal of the Statistical and Social Inquiry Society of Ireland,* XXXVIII.

VEST, J. & GAMM, L. 2010. Health information exchange: persistent challenges and new strategies. *JAMIA,* 17**,** 288-294.

WALKER, J., PAN, E., JOHNSTON, D., ADLER-MILSTEIN, J., BATES, D. W. & MIDDLETON, B. 2005. The Value Of Health Care Information Exchange And Interoperability. *Health Affairs*.

WANG, S., MIDDLETON, B., PROSSER, L., BARDON, C., SPURR, C., , P., KITTLER, A., GOLDSZER, R., FAIRCHILD, D., SUSSMAN, A., KUPERMAN, G. & BATES, D. 2003. A cost-benefit analysis of electronic medical records in primary care. *The American Journal of Medicine,* 114**,** 397-403.

WANNA, J. 2007. *Improving Implementation: Organisational Change and Project Management,* Canberra, WK Hancock Library, The Australian National University.

WHO. 2013. *National cancer control programmes* [Online]. World Health Organisation. Available: http://www.who.int/cancer/nccp/en/ [Accessed 16 Feb 2013].

WILKINSON, D., MICHIE, S. & MCCARTHY, M. 2007. The use and perceptions of routine health data: a qualitative study of four cancer network teams in England. *Health Serv Manage Res,* 20**,** 211-9.

WILKINSON, D. L. & MCCARTHY, M. 2007. Use of comparative data for integrated cancer services. *BMC Health Serv Res,* 7**,** 204.

WINSLOW, C. E. A. 1920. The Untilled Fields of Public Health. *Science,* 51**,** 23-33.

YIN, R. K. 2009. *Case Study Research: Design and Methods (4th Edition),* Thousand Oaks, CA, Sage.

Appendix 1. Ethics approval

From: Research Ethics <research-ethics@scss.tcd.ie>

Date: 18 December 2012 11:18

Subject: RE: research ethics application - Patricia Heckmann MSc HI

To: Patricia Heckmann <heckmanp@tcd.ie>

Cc: Lucy Hederman <hederman@scss.tcd.ie>, Research Ethics <research-ethics@scss.tcd.ie>

Dear Patricia,

Thank you for these revisions. You may now proceed with this study.

We wish you success in your research.

Kind regards

Gillian

Appendix 2. Gatekeeper approval

ncp National Cancer Control Programme

Feidhmeannacht na Seirbhíse Sláinte
Health Service Executive

Ms Patricia Heckmann
47 Edenvale Road
Ranelagh
Dublin 6

10th Jan 2013

Re: MSc HI - A Case Study to demonstrate the Contemporary Challenges to the Secondary Use of Health Data in Cancer Control in Ireland

Dear Ms Heckmann

I refer to your recent email and our meeting in relation to the proposed study and your request for information.

My Personal Assistant will supply you with the contact details of the appropriate people to contact in the PCRS and the NCCP, in addition to the details of the public hospitals involved in the provision of cancer services and also data with regard to hospital activity in the area of administration of systemic cancer treatment.

I will also organise that you have access to the test environment and the user manual of the Oncology Drug Management Scheme for the purpose of this study.

I would also like to take the opportunity to wish you well with your studies.

Yours sincerely

Susan O'Reilly

Dr. Susan O'Reilly MB, BCh, BAO, FRCPC, FRCPI
National Director
National Cancer Control Programme

Appendix 3. Informed consent form and information sheets

Title of research: A case study to demonstrate the challenges to the secondary use of health data in public health in Ireland.

Timeframe & duration of research: January – May 2013

DECLARATION:

- I have read, or had read to me, a document providing information about this research and this consent form. I have had the opportunity to ask questions, and all my questions have been answered to my satisfaction and I understand the description of the research that is being provided to me.
- I agree that my data is used for scientific purposes and I have no objection that my data is published in scientific publications or presentations in a way that does not reveal my identity.
- I understand that if I make illicit activities known, these will be reported, to appropriate authorities.
- I freely and voluntarily agree to be part of this research study, though without prejudice to my legal and ethical rights.
- I understand that I may refuse to answer any question and that I may withdraw at any time without penalty.
- I understand that any audio recordings will not be identifiable and will only be used for the purpose of making notes of the interview
- I understand that if I or anyone in my family has a history of epilepsy then I am proceeding at my own risk. (For those study phases involving computer monitors).
- I have received a copy of this agreement.
- I am 18 years or older and am competent to provide consent.

PARTICIPANT'S NAME: (please print)

PARTICIPANT'S SIGNATURE:

Date:

Statement of investigator's responsibility: I have explained the nature and purpose of this research study, the procedures to be undertaken and any risks that may be involved. I have offered to answer any questions and have fully answered such questions. I believe that the participant understands my explanation and has freely given informed consent.

RESEARCHER'S CONTACT DETAILS:
Patricia Heckmann
Patricia.Heckmann@cancercontrol.ie
Tel: 01 828 7184

RESEARCHER'S SIGNATURE:

Date:

Information Sheet for Participants - hospitals

Dear Participant,

I would like to invite you to take part in a research study entitled "A case study to demonstrate the challenges to the secondary use of health data in public health in Ireland”. This research study is being undertaken in part fulfilment of an MSc in Health Informatics in conjunction with the University of Dublin, Trinity College, Ireland.

Please read the following information carefully and please ask if you do not understand any part of it or would like more information.

Voluntary Participation

Your participation in this study is voluntary and you are free to withdraw at any time without providing a reason. If you do not wish to answer any specific questions, these wishes will be respected by the researcher.

Conflict of interest

The researcher has no conflict of interest to declare.

Name of Researcher: Patricia Heckmann

Timeframe & duration of research: January – May 2013

What is the purpose of the research study?

The purpose of the study is to establish the challenges to the secondary use of health data in public health in Ireland. The study is utilising the recently implemented PCRS reimbursement web portal as a case study. This portal requires manual entry of a minimum dataset as established by the NCCP/PCRS. However, it is possible that these data are already available in the hospital’s chemotherapy compounding software and that they could potentially be leveraged in order to reduce the requirement for manual data entry through electronic data transmission.

Why have you been chosen?

You have been chosen to participate in this study in your role as a provider of chemotherapy compounding software to Irish hospitals.

Who is organising the research study?

This study is being organised by the lead researcher Patricia Heckmann. There are no external collaborators involved in this study. The study is being supervised by a Trinity College supervisor. No funding is being provided for this study. The study will be completed between January 2013 and June 2013.

What will happen to me if I take part?

If you choose to take part in this study I will contact you to complete a questionnaire. The time taken to complete the questionnaire is anticipated to be approximately 20 minutes. Informed consent will be requested for the study.

Confidentiality - who will know I am taking part in the research study?

All information which is collected during the course of the research will be kept strictly confidential. The declaration of consent is integral to the questionnaire. The data collected will be aggregated for the purpose of the research and no participants or institutions will be individually named in my dissertation or any subsequent publications or presentations.

Personal data, save for purely communications purposes – e.g. telephone numbers and e-mail, will not be retained within the meaning of the Data Protection Act. This data will be held secure for the period of time required by the college. It will not be disseminated for any other purpose or be further processed in any other way.

How will data be stored and protected?

Data collection, storage and analysis will be in line with the Data Protection (& Amendment) Acts and Best Practice in Scientific Research. Individual results will be aggregated anonymously and research will be reported on aggregate results. No individual patient data will be collected for the purpose of this study.

Research Ethics Approval

The Research Ethics Committee of the School of Computer Science & Statistics, University of Dublin, Trinity College granted ethical approval for this study in December 2012.

What will happen to the results of the research study?

The results of the study will be presented in my dissertation for submission to the University of Dublin, Trinity College, and may be used by others for academic research. It is also proposed to submit the results to a peer-reviewed journal. In addition the research outcomes are likely to be presented at selected conferences, seminars or workshops.

The results will be made available, if requested, by email to all research participants on completion of the dissertation.

Procedure to be used if assistance or advice is needed

In the event that you require further information, assistance or advice about this study please contact Patricia Heckmann by email: Patricia.Heckmann@cancercontrol.ie or by phone: 01 8287184 and I will be happy to answer your questions.

If debriefing is required, it will be provided by the researcher within 4 weeks of completion of the questionnaire.

Thank you for taking the time to read this correspondence and for considering taking part in the research study.

Yours sincerely,

Ms. Patricia Heckmann

Information Sheet for Participants - interviews

Dear Participant,

I would like to invite you to take part in a research study entitled "A case study to demonstrate the challenges to the secondary use of health data in Cancer Control in Ireland”. This research study is being undertaken in part fulfilment of an MSc in Health Informatics in conjunction with the University of Dublin, Trinity College, Ireland.

Please read the following information carefully and please ask if you do not understand any part of it or would like more information.

Voluntary Participation

Your participation in this study is voluntary and you are free to withdraw at any time without providing a reason. If you do not wish to answer any specific questions, these wishes will be respected by the researcher.

Conflict of interest

The researcher has no conflict of interest to declare.

Name of Researcher: Patricia Heckmann

Timeframe & duration of research: January – May 2013

What is the purpose of the research study?

The purpose of the study is to establish the challenges to the secondary use of health data in cancer control in Ireland. The study is utilising the recently implemented PCRS reimbursement web portal as a case study. This portal requires manual entry of a minimum dataset (attached) to facilitate the reimbursement of the costs of systemic cancer therapy in hospitals.

Why have you been chosen?

You have been invited to participate in this study as a stakeholder in the domain of secondary data use and as a person positioned to inform the researcher as to the challenges to the use of this data.

Who is organising the research study?

This study is being organised by the lead researcher Patricia Heckmann. There are no external collaborators involved in this study. The study is being supervised by a Trinity College supervisor. No funding is being provided for this study. The study will be completed between January 2013 and May 2013.

What will happen if I take part?

If you choose to take part in this study, I will contact you to arrange a time to conduct a semi-structured interview. The interview will take place over the telephone, or face to face and will be of a duration of approximately thirty minutes. A series of lead questions have been prepared and will be supplied at least one week in advance of the arranged time to allow questions or clarifications to be dealt with in advance of the interview. The intention is to audio record the interview and also to make notes. If you do not wish to be recorded I will only take notes during the interview. Informed consent will be requested for the study.

Confidentiality - who will know I am taking part in the research study?

All information which is collected during the course of the research will be kept strictly confidential. The declaration of consent is integral to the questionnaire. The data collected will be aggregated for the purpose of the research and no participants or institutions will be individually named in my dissertation or any subsequent publications or presentations.

Personal data, save for purely communications purposes – e.g. telephone numbers and e-mail, will not be retained within the meaning of the Data Protection Act. This data will be held secure for the period of time required by the college. It will not be disseminated for any other purpose or be further processed in any other way.

How will data be stored and protected?

Data collection, storage and analysis will be in line with the Data Protection (& Amendment) Acts and Best Practice in Scientific Research. Individual results will be aggregated anonymously and research will be reported on aggregate results. No individual patient data will be collected for the purpose of this study.

Research Ethics Approval

The Research Ethics Committee of the School of Computer Science & Statistics, University of Dublin, Trinity College granted ethical approval for this study in December 2012.

What will happen to the results of the research study?

The results of the study will be presented in my dissertation for submission to the University of Dublin, Trinity College, and may be used by others for academic research. It is also proposed to submit the results to a peer-reviewed journal. In addition, the research outcomes are likely to be presented at selected conferences, seminars or workshops.

The results will be made available, if requested, by email to all research participants on completion of the dissertation.

Procedure to be used if assistance or advice is needed

In the event that you require further information, assistance or advice about this study please contact me by email (heckmanp@tcd.ie) or by phone (087 7690161) and I will be happy to answer your questions. If debriefing is required, it will be provided by the researcher within 4 weeks of completion of the interview.

Thank you for taking the time to read this correspondence and for considering taking part in the research study.

Yours sincerely,

Ms. Patricia Heckmann

Information Sheet for Participants - time study

Dear Participant,

I would like to invite you to take part in a research study entitled "A case study to demonstrate the challenges to the secondary use of health data in Cancer Control in Ireland”. This research study is being undertaken in part fulfilment of an MSc in Health Informatics in conjunction with the University of Dublin, Trinity College, Ireland.

Please read the following information carefully and please ask if you do not understand any part of it or would like more information.

Voluntary Participation

Your participation in this study is voluntary and you are free to withdraw at any time without providing a reason. If you do not wish to answer any specific questions, these wishes will be respected by the researcher.

Conflict of interest

The researcher has no conflict of interest to declare.

Name of Researcher: Patricia Heckmann

Timeframe & duration of research: January – May 2013

What is the purpose of the research study?

The purpose of the study is to establish the challenges to the secondary use of health data in cancer control in Ireland. The study is utilising the recently implemented PCRS reimbursement web portal as a case study. This portal requires manual entry of a minimum dataset (attached) to facilitate the reimbursement of the costs of systemic cancer therapy in hospitals.

Why have you been chosen?

You have been invited to participate in this study as you have some experience in the use of the PCRS web portal for the purpose of reimbursement of chemotherapy.

Who is organising the research study?

This study is being organised by the lead researcher Patricia Heckmann. There are no external collaborators involved in this study. The study is being supervised by a Trinity College supervisor. No funding is being provided for this study. The study will be completed between January 2013 and May 2013.

What is my role in this study?

If you choose to take part in this study I will contact you to arrange a time where I would observe the time required, by you, to input the required data into the PCRS web portal. The data input would utilise predefined scenarios, with patients of the hospital where available. No new patients or fictional patient history would be added to the PCRS file during the time study. The researcher will not observe the input screen so as to preserve patient confidentiality. The time taken to complete the time study is anticipated to be approximately 20 minutes. Informed consent will be requested for the study.

Confidentiality - who will know I am taking part in the research study?

All information which is collected during the course of the research will be kept strictly confidential. The declaration of consent is integral to the questionnaire. The data collected will be aggregated for the purpose of the research and no participants or institutions will be individually named in my dissertation or any subsequent publications or presentations.

Personal data, save for purely communications purposes – e.g. telephone numbers and e-mail, will not be retained within the meaning of the Data Protection Act. This data will be held secure for the period of time required by the college. It will not be disseminated for any other purpose or be further processed in any other way.

How will data be stored and protected?

Data collection, storage and analysis will be in line with the Data Protection (& Amendment) Acts and Best Practice in Scientific Research. Individual results will be aggregated anonymously and research will be reported on aggregate results. No individual patient data will be collected for the purpose of this study.

Research Ethics Approval.

The Research Ethics Committee of the School of Computer Science & Statistics, University of Dublin, Trinity College granted ethical approval for this study in December 2012.

What will happen to the results of the research study?

The results of the study will be presented in my dissertation for submission to the University of Dublin, Trinity College, and may be used by others for academic research. It is also proposed to submit the results to a peer-reviewed journal. In addition, the research outcomes are likely to be presented at selected conferences, seminars or workshops.

The results will be made available, if requested, by email to all research participants on completion of the dissertation.

Procedure to be used if assistance or advice is needed

In the event that you require further information, assistance or advice about this study please contact me by email (heckmanp@tcd.ie) or by phone (087 7690161) and I will be happy to answer your questions. If debriefing is required, it will be provided by the researcher within 4 weeks of completion of the interview.

Thank you for taking the time to read this correspondence and for considering taking part in the research study.

Yours sincerely,

Ms. Patricia Heckmann

Information sheet for participants – system vendors

Dear Participant,

I would like to invite you to take part in a research study entitled "A case study to demonstrate the challenges to the secondary use of health data in public health in Ireland”. This research study is being undertaken in part fulfilment of an MSc in Health Informatics in conjunction with the University of Dublin, Trinity College, Ireland.

Please read the following information carefully and please ask if you do not understand any part of it or would like more information.

Voluntary Participation

Your participation in this study is voluntary and you are free to withdraw at any time without providing a reason. If you do not wish to answer any specific questions, these wishes will be respected by the researcher.

Conflict of interest

The researcher has no conflict of interest to declare.

Name of Researcher: Patricia Heckmann

Timeframe & duration of research: January – May 2013

What is the purpose of the research study?

The purpose of the study is to establish the challenges to the secondary use of health data in public health in Ireland. The study is utilising the recently implemented PCRS reimbursement web portal as a case study. This portal requires manual entry of a minimum dataset as established by the NCCP/PCRS. However, it is possible that these data are already available in the hospital’s chemotherapy compounding software and that they could potentially be leveraged in order to reduce the requirement for manual data entry through electronic data transmission.

Why have you been chosen?

You have been chosen to participate in this study in your role as a provider of chemotherapy compounding software to Irish hospitals.

Who is organising the research study?

This study is being organised by the lead researcher Patricia Heckmann. There are no external collaborators involved in this study. The study is being supervised by a Trinity College supervisor. No funding is being provided for this study. The study will be completed between January 2013 and June 2013.

What will happen to me if I take part?

If you choose to take part in this study I will contact you to complete a questionnaire. The time taken to complete the questionnaire is anticipated to be approximately 20 minutes. Informed consent will be requested for the study.

Confidentiality - who will know I am taking part in the research study?

All information which is collected during the course of the research will be kept strictly confidential. The declaration of consent is integral to the questionnaire. The data collected will be aggregated for the purpose of the research and no participants or institutions will be individually named in my dissertation or any subsequent publications or presentations.

Personal data, save for purely communications purposes – e.g. telephone numbers and e-mail, will not be retained within the meaning of the Data Protection Act. This data will be held secure for the period of time required by the college. It will not be disseminated for any other purpose or be further processed in any other way.

How will data be stored and protected?

Data collection, storage and analysis will be in line with the Data Protection (& Amendment) Acts and Best Practice in Scientific Research. Individual results will be aggregated anonymously and research will be reported on aggregate results. No individual patient data will be collected for the purpose of this study.

Research Ethics Approval

The Research Ethics Committee of the School of Computer Science & Statistics, University of Dublin, Trinity College granted ethical approval for this study in December 2012.

What will happen to the results of the research study?

The results of the study will be presented in my dissertation for submission to the University of Dublin, Trinity College, and may be used by others for academic research. It is also proposed to submit the results to a peer-reviewed journal. In addition the research outcomes are likely to be presented at selected conferences, seminars or workshops.

The results will be made available, if requested, by email to all research participants on completion of the dissertation.

Procedure to be used if assistance or advice is needed

In the event that you require further information, assistance or advice about this study please contact Patricia Heckmann by email: Patricia.Heckmann@cancercontrol.ie or by phone: 01 8287184 and I will be happy to answer your questions.

If debriefing is required, it will be provided by the researcher within 4 weeks of completion of the questionnaire.

Thank you for taking the time to read this correspondence and for considering taking part in the research study.

Yours sincerely,

Ms. Patricia Heckmann

Appendix 4. Minimum data set – PCRS reimbursement portal

Section	Status*	Details
Demographics	R	Patient Surname
	R	Patient Forename
	R	Patient Date of Birth
	R	Patient Gender
	O	Patient Home Phone Number Patient Mobile Phone Number Patient email
	R	Patient Address, Line 1 - any text
	R	Patient Address, Line 2 - Town
	R	Patient Address, Line 3 - County
	R	Patient Address, Line 4 - Country
	R	Patient Identification number PPSN or Patient Identification number Medical Card or Patient Identification number DPS or Patient Identification number DV
	R	Patient Identification number type PPSN or Patient Identification number type Medical Card or Patient Identification number type DPS or Patient Identification number type DV
Diagnosis - ICD10 code	R	Drop down menu
Treatment protocol	R	Drop down menu
Dispensed Date of chemotherapy	R	Enter date – calendar option available
Drug dispensed	R	Drop down menu
Quantity of units	R	Number of vials, ampoules etc. dispensed

***Required (R) or Optional (O)**

Appendix 5. Data field comparison

The table below shows the data fields that are required to be completed in the NCCP/PCRS reimbursement system. These are compared to data fields from other national data collections. Where "own data dictionary" is mentioned it indicates that the data dictionary differs from that of the NCCP/PCRS reimbursement scheme data dictionary, am example of this is the county code fields as detailed in 0.

Section	Status*	Details	NCRI	HIPE	PCRS**	NCCP KPIs
Demographics	R	Patient Surname	Yes	No	Yes	No
	R	Patient Forename	Yes	No	Yes	No
	R	Patient Date of Birth	Yes	Yes	Yes	No
	R	Patient Gender	Yes – own data dictionary	Yes	??	No
	R	Patient Address, Line 1		No	Yes	No
	R	Patient Address, Line 2- Town		No	Yes	No
	R	Patient Address, Line 3 - County	Yes – own data dictionary	Yes – own data dictionary	Yes – own data dictionary	No
	R	Patient Address, Line 4 - Country	Yes – own data dictionary	Coded area of residence – own data dictionary	Yes – own data dictionary	No
	R	Patient Identification number PPSN	Yes	No	Yes	No
		or Patient Identification number Medical Card or Patient Identification number DPS or Patient Identification number DV	Yes Medical Card Identifier	Yes - Patient Identification number Medical Card	Yes	No
	R	Patient Identification number type PPSN or Patient Identification number type Medical Card or Patient Identification number type DPS or Patient Identification number type DV	No	Yes – medical Card Status	Yes	No

Section	**Status***	**Details**	**NCRI**	**HIPE**	**PCRS****	**NCCP KPIs**
Diagnosis - ICD10 code	R	Drop down menu	Yes – own data dictionary	Yes – own data dictionary	No	No
Treatment protocol	R	Drop down menu	Yes	Yes	No	
Dispensed Date of chemotherapy	R	Enter date – calendar option available	Yes - own data dictionary		Yes	Date of first treatment
Drug dispensed	R	PCRS code - Drop down menu	No	No	Yes	No
Quantity of units	R	Number of vials, ampoules etc. dispensed			Yes	No
Fields not in the selected case						
MRN			Yes	Yes	No	No
Treating medical Consultant			Yes	Yes	No	No
Hospital Code			Yes	Yes	No	No
Health insurer			No	Yes	No	No
Oncology day ward flag – chemotherapy initial and repeat encounters			No	Yes	No	No
		(NCCP, 2012a)	(NCRI, 2013)	(HIPE, 2013)		(NCCP, 2013b)

*Status R = required field

**Community drugs reimbursement

NCCP KPIs relate to medical oncology KPIs only

Appendix 6. County code data dictionaries

NCRI data dictionary - County

NCRI county_id	description
CE	CLARE
CK	CORK
CN	CAVAN
CW	CARLOW
DL	DONEGAL
DN	DUBLIN
GY	GALWAY
KE	KILDARE
KK	KILKENNY
KY	KERRY
LD	LONGFORD
LH	LOUTH
LK	LIMERICK
LM	LEITRIM
LS	LAOIS
MH	MEATH
MN	MONAGHAN
MO	MAYO
OS	OVERSEAS/NI/GB/UK
OY	OFFALY
RN	ROSCOMMON
SO	SLIGO
TY	TIPPERARY
WD	WATERFORD
WH	WESTMEATH
WW	WICKLOW
WX	WEXFORD
ZZ	UNKNOWN

County code data dictionary - ODMS

Antrim	Donegal	Dublin 6	Dublin 12	Dublin 20	Kildare	Mayo	Tyrone
Armagh	Down	Dublin 6W	Dublin 13	Dublin 22	Kilkenny	Meath	Waterford
Carlow	Dublin 1	Dublin 7	Dublin 14	Dublin 24	Laois	Monaghan	Westmeath
Cavan	Dublin 2	Dublin 8	Dublin 15	Dublin	Leitrim	Offaly	Wexford
Clare	Dublin 3	Dublin 9	Dublin 16	Fermanagh	Limerick	Roscommon	Wicklow
Cork	Dublin 4	Dublin 10	Dublin 17	Galway	Longford	Sligo	
Derry	Dublin 5	Dublin 11	Dublin 18	Kerry	Louth	Tipperary	

RESID – Area of Residence (HSE Region only applies to Republic of Ireland)				
Code	**Area of Residence**	**HSE Region**	**Code**	**Area of Residence**
0000	No fixed abode	N/A	3310	Northern Ireland
0500	Carlow	HSE South	3318	Austria
2900	Cavan	HSE Dublin North East	3319	Switzerland
1600	Clare	HSE West	3301	Belgium
1101	Cork City	HSE South	3331	Bulgaria
1200	Cork County	HSE South	3321	Cyprus
2800	Donegal	HSE West	3322	Czech Republic
0100	North Dublin City and County	HSE Dublin North East	3302	Denmark
0200	South Dublin City and County	HSE Dublin Mid-Leinster	3303	England
1801	Galway (City)	HSE West	3323	Estonia
1900	Galway (County)	HSE West	3316	Finland
1300	Kerry	HSE South	3304	France
0300	Kildare	HSE Dublin Mid-Leinster	3305	Germany
0700	Kilkenny	HSE South	3306	Greece
2500	Laois	HSE Dublin Mid-Leinster	3307	Netherlands
2600	Leitrim	HSE West	3324	Hungary
1401	Limerick City	HSE West	3308	Italy
1500	Limerick County	HSE West	3325	Latvia
2200	Longford	HSE Dublin Mid-Leinster	3326	Lithuania
3100	Louth	HSE Dublin North East	3309	Luxembourg
2100	Mayo	HSE West	3327	Malta
3200	Meath	HSE Dublin North East	3317	Norway
3000	Monaghan	HSE Dublin North East	3328	Poland
2400	Offaly	HSE Dublin Mid-Leinster	3314	Portugal
2000	Roscommon	HSE West	3332	Romania
2700	Sligo	HSE West	3311	Scotland
1700	Tipperary North Riding	HSE West	3329	Slovak republic
0800	Tipperary South Riding	HSE South	3330	Slovenia
0901	Waterford City	HSE South	3313	Spain
1000	Waterford County	HSE South	3315	Sweden
2300	Westmeath	HSE Dublin Mid-Leinster	3320	Wales
0600	Wexford	HSE South	3350	Other European
0400	Wicklow	HSE Dublin Mid-Leinster	3400	Africa
3501			America North and Canada	
3502			America South, Central and Caribbean	
3600			Asia, Middle East, Far East, South East Asia	
3700			Australia, New Zealand, Oceania	

Appendix 7. Data collection tool - time study

Preamble: As per our earlier correspondence you will be aware that I am a student currently undertaking a Trinity College Dublin Masters in Health Informatics. As a component of my Masters I am undertaking a research study. The purpose of the study is to demonstrate the challenges to the secondary use of health data in cancer control in Ireland. The study is utilising a recently implemented reimbursement web portal as a case study. The portal requires manual submission of data to facilitate reimbursement for systemic anti-cancer therapy (SACT) administered in public hospitals.

The purpose of this phase of the study is to determine the length of time taken to input data into PCRS web portal for the purpose of reimbursement for those SACT administered to patients in the hospital.

Time study participant:

Time study venue:

Date:``Time:

Consent received: Written Verbal

Data reporting work elements:

The process of data entry can be broken into five discrete parts.

1. Login
2. New patients – i.e. cycle one of chemotherapy (limitation some patients would have being potentially treated for other cancers previously, may be a recurrence, relapse)
    a. Patient registration – existing patient on PCRS file – add three patients
    b. Patient registration – patient not on PCRS file – add three patients
3. Finding a patient – patients already registered – find three patients from above
4. Changing a patient's diagnosis – change the diagnosis of the three patients from point 3.
5. Adding drugs utilised in the treatment of the patient – add one line of drug data for three patients and two lines of drug data for three patients (may be the same patients)

Limitations: - does not include time to collate data or time to follow up on payments of entries etc.

Test patient information for use in test system. X = interviewee number

New patient entry not on PCRS

Patient Forename Name	Test	Test	Test
Patient Surname	TestXa	TestXb	TestXc
Patient address	Xa Test Street, Test Town Dublin 1 Ireland	Xb Test Street, Test Town Dublin 1 Ireland	Xc Test Street, Test Town Dublin 1 Ireland
Gender	Male	Female	Male
Patient DOB	x/x/1966	x/x/1966	x/x/1966
Protocol	Not assigned	Not assigned	Not assigned
Diagnosis ICD10	C76	C76	C76
PPSN	1234567T	1234567T	1234567T

Notes:

1. Start time = the user's first orientation toward the application following explanation and discussion of the scenario.
2. End time = the point when the user signals verbally that the task is completed.
3. Login = measure from the point where the user selected the URL for the portal. Time taken to find the link, start computer etc was not timed.
4. Patient registration / location = measure from the point of selection of the patient search/registration screen

Time study data collection tool	**Time required (min – sec)**	**Comments**
Login 1		
Login 2		
Login 3		
New patient registration 1 – on PCRS (live only)		
New patient registration 2 – on PCRS		
New patient registration 3 – on PCRS		
New patient registration 1 – not on PCRS		
New patient registration 2 – not on PCRS		
New patient registration 3 – not on PCRS		
Locating a patient (existing patient) 1		
Changing a patient diagnosis 1		
Locating a patient (existing patient) 2		
Changing a patient diagnosis 2		
Locating a patient (existing patient) 3		
Changing a patient diagnosis 3		
Add a drug line 1		
Add a drug line 1		
Add a drug line 1		
Other steps		
Other steps		
Other steps		

Appendix 8. Coded thematic analysis

Summary of thematic analysis (Detail in Section xx.xx): The researcher created individual word documents for each interview using the interview topics as a high-level thematic framework to focus the analysis. Interview notes were verified by the interviewees. Data analysis was manual, utilising word documents and spreadsheets. The data was initially categorised into high-level themes and then copied into spreadsheet rows; each column dedicated to a unique interview. Data was replicated where it was found to be consistent with more than one theme. Subcategories identified through the literature review were utilised as preset categories and supplemented by additional subcategories as they emerged. Additional themes were added by the researcher as they became apparent. This process continued until all data had been documented and analysed to the point where no new themes or subcategories were emerging. At this point the themes and subcategories were examined for duplicates and as a result the number of themes and subcategories was reduced from 156 to 117.

	Interview code	**2**	**5**	**9**	**11**	**17**
	Stakeholder group	**Hospital user**	**NCCP**	**HIQA**	**HSE ICT**	**Hospital user**
	Primary or Secondary Stakeholder	**Primary stakeholder**	**Primary stakeholder**	**Secondary stakeholder**	**Secondary stakeholder**	**Secondary stakeholder**
	Subset of stakeholder groups	**Pharmacist – cancer centre**	**Medical Oncology Programme**	**Standards**	**HSE ICT**	**Interface/ EHR specialist**
	Phone interview	**N**	**N**	**N**	**N**	**N**
	Interview recorded	**Y**	**N**	**Y**	**Y**	**Y**
	Notes Sent	**15/05/2013**	**16/05/2013**	**27/05/2013**	**27/05/2013**	**27/05/2013**
	Notes Verified	**20/05/2013**	**16/05/2013**	**27/05/2013**		
	Notes updated	**20/05/2013**	**16/05/2013**	**27/05/2013**		
Section number	**Theme**	**Notes**	**Notes**	**Theme**	**Theme**	**Theme**
1	**1. Benefits of collecting data**	**1. Benefits of collecting data**	**1. Benefits of collecting data**	**1. Benefits of collecting data**	**1. Benefits of collecting data**	**1. Benefits of collecting data**
Comments				Primary purpose of this data collection would be to enable delivery of care to the patient. "The data should be collected as a direct output of care (primary purpose) and used for secondary	""Not just about cancer. The more information and knowledge we have to support decision making; the better. This is all about decision making at the end of the day" That includes cancer diagnosis, treatment, money. "The more information	To detect incidence, treatment and prevention.

	Interview code	**2**	**5**	**9**	**11**	**17**
	Stakeholder group	**Hospital user**	**NCCP**	**HIQA**	**HSE ICT**	**Hospital user**
				purposes rather than a separate system to collect the data for secondary use" "Should only collect the data once and reuse it".	the better" . Concerned with the way we collect data. "Very focussed on collecting data. Invest an inordinate amount of time and energy in collecting data and we don't think how we could do it differently, as collecting is part of a process and should be a by-product of a process". "We should look at the way we do things and see how we can derive data as a by-product of that, particularly technology. " "Because we are so behind in understanding ICT supporting process that we invest in data collection and try and manipulate it and deal with it. This is quite a costly exercise that could be done differently"	
1a1	Service planning - national requirements, trends in incidence, local requirements	National Service Planning requirements. Local service planning requirements	National Service Planning requirements	Service planning - national requirements. Service planning - local requirements		
1a2	Service planning - financial	Financial trends - budgeting, benchmarking				
1b	Strategy implementation - health reform					
1c1	Audit - value/effectiveness/outcomes and impact of new drugs/ services	Audit of practice changes / new drugs		Audit - value/effectiveness/improve outcomes/research. National and international statistics will allow bench marking against other hospitals and other countries.		
1c2	Audit - adherence to national standards/guidelines/best practice, impact of practice variations	Audit impact of different practices between centres/bench marking	Audit of adherence to national guidelines/protocols			
1c3	Audit - performance management - KPIs		Audit/performance management - quantity, volume, timeliness of treatment, data for KPIs			
1d	Clinical trials/research					
1x	**Miscellaneous**					

	Interview code	**2**	**5**	**9**	**11**	**17**
	Stakeholder group	**Hospital user**	**NCCP**	**HIQA**	**HSE ICT**	**Hospital user**
2	**2. Agencies that collect data on cancer or for cancer control**	**2. Agencies that collect data on cancer or for cancer control**	**2. Agencies that collect data on cancer or for cancer control**	**2. Agencies that collect data on cancer or for cancer control**	**2. Agencies that collect data on cancer or for cancer control**	**2. Agencies that collect data on cancer or for cancer control**
		NCCP, HIPE, NCRI, HSE (financial and general data)	NCCP/PCRS, ESRI, HIPE, Irish cancer society, NCRI	NCRI, HIPE (ESRI), HIQA (audit)	NCRI	NCRI, ESRI (HIPE), ICS (research) , NTPF potentially as part of the patient treatment register (PTR) - episode specific for waiting lists, attendances, diagnosis sent for all patients. PTR wants a lot of new data
3	**3. Mandatory vs. Optional reporting**	**3. Mandatory vs. Optional reporting**	**3. Mandatory vs. Optional reporting**	**3. Mandatory vs. Optional reporting**	**3. Mandatory vs. Optional reporting**	**3. Mandatory vs. Optional reporting**
		NCRI	NCRI - not mandatory but allowed to collect the data under an exemption in the data protection legislation			ESRI and NCRI - mandatory
4	**4. Standards used in current data collection**	**4. Standards used in current data collection**	**4. Standards used in current data collection**	**4. Standards used in current data collection**	**4. Standards used in current data collection**	**4. Standards used in current data collection**
		Unsure, minimum datasets for NCRI to be collected	Data protection legislation	ICD10 mostly	No specific knowledge. There are minimum standards in terms of coding etc. Coding should e used where possible "If we are going to collect data at least we should collected it in a structured format. If we don't collect it in a structured format it will be difficult to get any real benefit from it". There is a lack of understanding wrt the requirement for structured data. The HSE is working on a project the "Integrated Services Framework (ISF)". Key to this is the development and adoption of an Irish data model for health in that would be standards for coding for diagnosis, communications, treatment, identity management, process, drugs etc. All the underlying architecture for systems and going forward all systems will have to comply with these" . Some of these exist but it would be to decided which ones we will use in Ireland. There is also a group in the SDU collecting data sets in use nationally. "Lots of groups doing lots of	NTPF group put together the requirements of what is sent, it uses standards, is coded and has a data dictionary. HIPE codes were reused, IDC03, ICD10,, the NCRI dataset uses the same fields as HIPE as well. Codes for areas if residence, discharge type, consultants.

	Interview code	2	5	9	11	17
	Stakeholder group	**Hospital user**	**NCCP**	**HIQA**	**HSE ICT**	**Hospital user**
					things with little oversight over the complete picture". "The HSE has complied an inventory of hospital systems that are currently being supported in the HSE; it currently stands at 1,752 systems in use, different versions of the same system, this would include access databases containing clinical information". "This results in data in silos, non validated non structured". Trying to pull it together by agreeing data standards that would be used in systems across the HSE. The clinical care programmes, each wants its own data set and a system to support them. Even the basic demographic dataset have not been agreed. "HIQA have a draft national demographic data set, the problem is in getting legacy systems to comply with this data set and to map existing demographic fields to the new standards, which is still draft". "There is an enormous amount of data around but I would just question its usefulness; for the specific purpose for which it was collected probably yes; but in terms of its real use to the whole health system - questionable". "It is time for the big picture to be seen and acted on and for people and institutions to stop doing their own thing" "There is no need to reinvent wheels all the time". "In the health industry there are effectively thousands of small companies providing a service, every GP and consultant is a private company. They run the business they want to run the business."	
5	**5. Challenges to the secondary use of data**	**5. Challenges to the secondary use of data**	**5. Challenges to the secondary use of data**	**5. Challenges to the secondary use of data**	**5. Challenges to the secondary use of data**	**5. Challenges to the secondary use of data**
5	**Priority volunteered**				Variations in measurement system	
5	**Comment**			Secondary use - issues with regard to consent privacy and confidentiality. Data protection legislation can be problematic in	All the themes raised are issues and challenges but something can be done about them as can be done with regard to variations in measurement systems. The	"Needs a national lead or legislation to dictate the standards to be used so that there will be consistency"

	Interview code	**2**	**5**	**9**	**11**	**17**
	Stakeholder group	**Hospital user**	**NCCP**	**HIQA**	**HSE ICT**	**Hospital user**
				some areas unless the patient is deidentified, but then you cannot track your patient. "Lack of unique identifiers is a problem with patients receiving services in different places, need to be able to track the patient through the care pathway. Identifiers would be of huge benefit for patients, organisations and practitioners".	problem is with all the data we have already which is good data and the costly effort once we decide on standards too integrate it into the new structure.	
5a1	Inefficiency - documentation and data quality - duplication of work					
5a2	Inefficiency - staffing resources					
5a3	Inefficiency - documentation and data quality	Collating/aggregating the data. Hospitals not ready to give the data - no aggregation in place				"To use data meaningfully you need to have consistent and quality data to start with. "
5a4	Inefficiency - documentation and data quality - data quality, completeness and quality assurance	Quality of data / data inaccuracies / incomplete data / invalidates the entire data set.	Data quality - no QA for example through mechanisms such as duplicate data entry for verification of record accuracy, timeliness, accuracy	The data quality is an issue, quality in all its different dimensions, including legibility, completeness, timeliness, and so on		
5a6	Inefficiency - documentation and data quality - local data collation/aggregation	Data quality - errors		"An army of people (coders) that deal with HIPE data, and other army of people dealing with cancer, another army that deal with this or the other...an awful lot of duplication and fragmentation, variable quality". "Not only is this inefficient and results in increased cost but also mitigates against reusing the data, maximising the reusability of the data." "It does not matter what area you are talking about data is expensive to collect and you really want to make as much use f it as you can		

	Interview code	**2**	**5**	**9**	**11**	**17**
	Stakeholder group	**Hospital user**	**NCCP**	**HIQA**	**HSE ICT**	**Hospital user**
				for secondary purposes as well as primary purposes" . "The Irish health system does not recognise the profession of people who are responsible for data quality and so on, as a result ensuing the quality of the data is everybody's responsibility and nobodies, there is not clear accountability and responsibility for the data. This is especially pertinent for secondary use as it is dependent on the data collection and quality implemented for primary use."		
5a, 5d1	Inefficiency, Technological barriers - location of data - paper charts/systems/not recorded	Location of data - paper charts vs. electronic, different electronic systems				
5a, 5d2	Inefficiency, Technological barriers - disparate/legacy systems	Disparate electronic and legacy systems recording patient data.	Disparate electronic and legacy systems recording patient data.	"Tremendous fragmentation and silos"		
5b1	Variations in measurement systems - national	Different data reporting requirements for different stakeholders				There is no national systems for data collections, lots of little national systems to report to, they dictate how the data is reported. A lot of them are similar and some have collaborated, HIPE, NCRI, but all can set their own standards and data dictionaries. There is some reuse of fields and coding standards for example ICD10 is the same in each system. HIQA are now developing a demographics dataset and the NCRI have one already. "One reasons for inefficiency is that the same data is required by many reporting systems, mainly the same but some have variations. This results in duplicate data collection and reporting, whether manual or electronic." Also have reporting to insurance

	Interview code	**2**	**5**	**9**	**11**	**17**
	Stakeholder group	**Hospital user**	**NCCP**	**HIQA**	**HSE ICT**	**Hospital user**
						companies
5b2	Variations in measurement systems - local					
5b3	Variations in measurement systems - mandatory vs. Voluntary					"If the reporting of data is to be mandatory there needs to be penalties like those financial penalties in place for exceeding waiting list times".
5c1	Organisational and cultural issues - work practice					
5c2	Organisational and cultural issues - not my job					
5c3	Organisation and cultural issues - training					
5c4	Organisation and cultural issues - cooperation		Culture - willingness of people to cooperate or not			
5c5	Organisation and cultural issues - human barriers					
5c6	Organisation and cultural issues - eligibility of treatment					
5c7	Organisation and cultural issues - keeping staff up to date					
5d1	Technological barriers - general comment				"There is no technological barrier to the EHR" . In 2005 the RAND corporations they did a study looking the economic value of the EHR. In 2012 they reviewed what happened and the two things they found interoperability and ease of use but thirdly and the main one - the health industry did not look at the way they did their business process. "Typically, in health, we bring in a system simply to replicate what someone is actually doing and not look at how they do it, why they do it, and what value is it. That's what we	

	Interview code	**2**	**5**	**9**	**11**	**17**
	Stakeholder group	**Hospital user**	**NCCP**	**HIQA**	**HSE ICT**	**Hospital user**
					haven't done and that's the barrier, it's not a technology barrier. It is process reengineering ". "The technology works, may not be easy to use". "The real barrier is that as an industry we have not changed the way we do our business to take advantage of new technologies" . There are fears that the introduction of technology will bring about redundancies. Why can we not make appointments in the hospitals ourselves. "Half hearted implementation, lack of decision making, fear of change" this is all half hearted implementations. "System works" - airline industry 90% of accidents are caused by human error, pilot intervention or lack of it cause crashed, it would be safer statically not to let a pilot fly a plane. Mist incidences in hospitals are caused by humans, take the humans out = safer system. "People that is the barrier" . Green field technology implementations are easier as up to date and can maximise the use of technology.	
5d2	Technological barriers - lack of skill					
5d3	Technological barriers - data sets	Lack of data sets				There is a need for consistency in naming conventions, the use of codes and data dictionaries. Inconsistencies lead to inaccurate data and duplication of records. National Client Index (NCI) - duplication of patients.
5d4	Technological barriers - security, privacy, ownership		System security - are they up to date?			PPSN is not to be used without consent yet is key to the PCRS dataset for patient identification. It is becoming a defacto national identifier. There are data protection issues when reporting data and the requirement for patient consent.
5d5	Technological barriers - uncoordinated implementation					

	Interview code	**2**	**5**	**9**	**11**	**17**
	Stakeholder group	**Hospital user**	**NCCP**	**HIQA**	**HSE ICT**	**Hospital user**
5d6	Technological barriers - interoperability	Lack of system interoperability / communication / prevents data reuse - within and between institutions. No standardisation				
5d7	Technological barriers - identifiers	Lack of Unique patient identifier	Lack of Unique patient identifier			There are some national systems e.g. NIMIS which also have a data set and dictionary but cannot find patients. There are a lot of patients duplicated in the systems as a result and it is episodic rather than patient centric. NTPF and HIPE are episode specific so the unique identifier is not important.
5d8	Technological barriers - not ready for task					
5d9	Technological barriers - cost / indecision					
5e1	Economic pressures - resources staff	Resources staffing	Resources staffing - primary collection of data is in the process of clinical care - the reporting always goes back to this point			
5e2	Economic pressures - resources equipment	Resources - equipment - hardware, software,	Resources - financial - cost of systems, interfaces, training, maintenance, upgrades, ongoing cost. No money available to fund these			"Technology can be sorted if there is money for resources"
5e3	Economic pressures - no incentives					
5e4	Economic pressures - communication and dissemination of information		Communication with all the relevant staff - particularly front line staff			
5f 1	Competing priorities - data sharing					
5f 2	Competing priorities -	Data protection,	Patient consent - if			

	Interview code	**2**	**5**	**9**	**11**	**17**
	Stakeholder group	**Hospital user**	**NCCP**	**HIQA**	**HSE ICT**	**Hospital user**
	consent/data protection legislation	anonymising data. Lack of legislation on use of patient data.	required by the DPC. Data protection - Henrietta Lacks - the classic example of unconsented retention of cells, and samples - later recognised -huge benefit to humanity. Guthrie card destruction with no risk benefit analysis on the effect on the individual or society.			
5f4	Competing priorities - local administrative vs. Local clinical priorities	Prioritisation of reporting - "It's always going to be bottom of the list of things to do as not seen as the core of anyone's work" "not direct patient care" so always last	Prioritisation - competing priorities with regard to collecting data and disseminating requirements and results			
5f5	Competing priorities - local vs. National	Prioritisation of local needs over national needs, prioritisation of all needs e.g. Some national requirements perceived as more beneficial or urgent to report than others. National data collection seen as "meddling by bureaucrats". Perception of no benefit locally with the provision of data.	Prioritisation of local needs over national needs, prioritisation of all needs e.g. Some national requirements perceived as more beneficial or urgent to report than others			
5x1	Miscellaneous - fear of misrepresentation or incorrect data used in reports					
5x2	Miscellaneous - reports as part of system specification?					
5x3	Miscellaneous - lack of consultation					

	Interview code	**2**	**5**	**9**	**11**	**17**
	Stakeholder group	**Hospital user**	**NCCP**	**HIQA**	**HSE ICT**	**Hospital user**
5x5	Miscellaneous					
6	**6. Why the SACT reimbursement scheme**	**6. Why the SACT reimbursement scheme**	**6. Why the SACT reimbursement scheme**	**6. Why the SACT reimbursement scheme**	**6. Why the SACT reimbursement scheme**	
6	**Comment**				The collection of this data is based on the absolute acceptance of the prescriber that this is the right course of action. It is a reimbursement system. This system is an example of a system to collect data only rather than a system that collect the data automatically as part of processes.	
6a1	Finance - budget		Consolidation/ centralisation of budget			
6a2	Finance - budget - money follows the patient			Finance - budget - money follows the patient		
6a3	Finance - budget - service cost / VFM	Budgeting and anticipating new drug uptake				
6a4	Finance - budget - allocation					
6a5	Finance - budget - accountable approved/formulary use only					
6b1	New data - prevalence / incidence of disease, etc	Data collection for statistical purposes down the line, patient numbers, prevalence of different diseases				
6b2	New data - cost negotiation	Additional data for NCCP when negotiating with agreed discounts with the drug companies				
6c1	Audit - clinical - adherence to protocols		Clinical audit - drugs being utilised for the approved indications			
6c2	Audit - clinical - outcomes		Outcome analysis			
6c3	Audit - clinical/VFM		Effectiveness research			
6c4	Improve service -	Standardise financial				

	Interview code	**2**	**5**	**9**	**11**	**17**
	Stakeholder group	**Hospital user**	**NCCP**	**HIQA**	**HSE ICT**	**Hospital user**
	outcomes, access, standardisation	arrangements for reimbursement of drugs - will benefit patients, no post code lottery				
6c5	Improve service - quality and safety					
6c5	Improve service - equitable access					
7	**7. Stakeholders in the SACT scheme**	**7. Stakeholders in the SACT scheme**	**7. Stakeholders in the SACT scheme**	**7. Stakeholders in the SACT scheme**	**7. Stakeholders in the SACT scheme**	**7. Stakeholders in the SACT scheme**
7	Stakeholders in the SACT scheme	NCCP, PCRS, DOH, HSE, (HSE/DOH/NCCP all the same organisation with regard to reporting)	Primary stakeholders - NCCP, PCRS (facilitating the process), Hospital staff (doctors, nurses, pharmacists, financial), Secondary stakeholders - patients, HSE/CPU, drug companies, NCCP/TRC			
7a	Groups with stake in new data				Pharmaceutical industry in looking at data in existing eHealth systems in terms of research. Trying to create a significant database. Big data exercise at European levels to create a specialist or complete expert system. The "virtual human being" - if an avatar was developed it could potentially circumvent the requirement for clinical trials. They experience all the same challenges.	
8	**8. Wider role for the data?**	**8. Wider role for the data?**	**8. Wider role for the data?**	**8. Wider role for the data?**	**8. Wider role for the data?**	**8. Wider role for the data?**
8a	Need to develop a new system for these data	Yes, a missed opportunity to build a database that would have more impact with regard to audit and research		These data are being collected for reimbursement primarily, so that is the focus of the data. Similarly with the HIPE system which was designed for casemix not quality indicators and outcome data collection. Clinicians were not involved and this is now beginning to change.		
8b	Cancer/Clinical information system	Integrate NCRI, HIPE, PCRS to create a	Data is very specific - hard to see the use	These data are a subset of data that could be produced from an		

	Interview code	**2**	**5**	**9**	**11**	**17**
	Stakeholder group	**Hospital user**	**NCCP**	**HIQA**	**HSE ICT**	**Hospital user**
		National Clinical/Cancer information system.	outside of the proposed current uses	electronic prescribing system. Also they are potentially a rich source for researchers.		
9	**9. Other data that could be collected in the SACT scheme**	**9. Other data that could be collected in the SACT scheme**	**9. Other data that could be collected in the SACT scheme**	**9. Other data that could be collected in the SACT scheme**	**9. Other data that could be collected in the SACT scheme**	**9. Other data that could be collected in the SACT scheme**
9a	New fields	Could capture a lot more data. So much could be added done to a micro level e.g. Within pharmacy, which diluents etc. Need to be pragmatic with regard to what is needed. Adjuvant vs. Neoadjuvant treatment, Tumour staging designations - would need national standardisation, Disease presentation - relapsed, refractory disease	Could capture a lot more data. So much could be added but how useful would it be. Need to be pragmatic with regard to what is needed	Insurance status - will be needed for the Universal Health insurance.	Consultants, adherence to protocols	Prescriber, area of residence (both have existing national codes in HIPE also a national speciality code)
9b	Comments on existing fields	ICD10 code in use is very high level		Where possible a minimum data set should only include those fields that are actually required. Coding should be used where possible and free text kept to a minimum. Drug dispensed will be a problem as there is no Irish catalogue. There are commercial catalogues such as the one distributed by the IPU and Helix Health	Rules would need to be put in place to authorise payment in the context of the protocol.	
10	**10. Missing electronic data (approximately 10%)**	**10. Missing electronic data (approximately 10%)**	**10. Missing electronic data (approximately 10%)**	**10. Missing electronic data (approximately 10%)**	**10. Missing electronic data (approximately 10%)**	**10. Missing electronic data (approximately 10%)**
10a		Transmit existing data and add missing data manually			Should extract the 90% and take it to another system. Link the data together using a flat field extract and use that data.	
10b			Have the missing data added to the existing data before transmission	If data is to be extracted and sent to another system then a standard message should be used for example HL7 and Healthlink as		2 step approach. Could look at the HIPE model. Hipe has hundreds of fields. The problem is patient identifier as HIPE is episode based

	Interview code	**2**	**5**	**9**	**11**	**17**
	Stakeholder group	**Hospital user**	**NCCP**	**HIQA**	**HSE ICT**	**Hospital user**
				this is in use by hospitals at present.		
10c		Change pharmacy systems to include the missing data if this is where the data extract is coming from	Change hospital systems to include the missing data if this is where the data extract is coming from	Electronic data should be extracted and reused if possible to minimise the workload and avoid transcription errors.		
11	**Main challenges to SACT scheme**	**11. Main challenges to SACT scheme**	**11. Main challenges to SACT scheme**	**11. Main challenges to SACT scheme**	**11. Main challenges to SACT scheme**	**11. Main challenges to SACT scheme**
11	**Comment**	All the challenges that pertain to the secondary use of data	All the challenges that pertain to the secondary use of data	All the challenges that pertain to the secondary use of data	As discussed in section 5	As discussed in section 5. Also manual entry, no identifier.
11a	Inefficiency - documentation and data quality - manual entry	Manual process - should be automated	Incomplete data, manual extraction, paper charts - data not available electronically within the hospital			
11b	Organisation and cultural issues - Cooperation between stakeholders		Cooperation between stakeholders - hospitals/NCCP/PCRS			
11c	Organisation and cultural issues - human barriers					
11x	Miscellaneous	Data not reported in real time	Agreement of the model - ownership			
12	**12 Facilitators**	**12 Facilitators**	**12 Facilitators**	**12 Facilitators**	**12 Facilitators**	**12 Facilitators**
12	**Comment**	Opposite to challenges	Opposite to challenges	Any system involving drugs needs standards and identifiers, patients, practitioners, doctors, pharmacists, drugs.	Opposite to challenges	Opposite to challenges
12a	Comment - legislation		Legislation with regard to personal health information, and in the interim a good working relationship with the DPC			
12b	Comment - motivation of staff		Support for staff - need to feel their work is valued and is a priority for the hospital -i.e. There is a value to their work			

	Interview code	**2**	**5**	**9**	**11**	**17**
	Stakeholder group	**Hospital user**	**NCCP**	**HIQA**	**HSE ICT**	**Hospital user**
12c	Comment - identifiers and interfaces		Unique identifiers for patients. Interfaces between systems			Identifiers for patients, providers, practitioners
12d	Comment - standard protocols for drug use		Standard protocols for drug use			
12e	Comment - public information/reassurance with regard to data use		Good communication to patients with regard to data sharing and consent - demonstrate the greater good principal and societal benefit. Demonstrating of good security, and methods of dealing with security breaches are required			
12f	Comment - better IT systems and systems that enable reuse of data - collect once and reuse	Efficient data capture. Automate the collation and reporting. Capture at a central location at admission/diagnosis.	Improved and consistent IT systems - or a single system			
12g	Stakeholder agreement and commitment		strong national lead			
12h	Comment resources / financial support					
12i	Comment - build on current data		Clear achievable plans for the data being collected			
13	**13. General Comments**	**13. General Comments**	**13. General Comments**	**13. General Comments**	**13. General Comments**	**13. General Comments**
13a	Comment					
13b	Comment - duplication of work			The collection of diagnosis manually will mean that at a minimum a patients diagnosis is being coded three times, HIPE, health insurance, this system, this is a very real demonstration of the duplication of work that these systems introduce which could be avoided by the "collect once and reuse" concept.		
13d	Pragmatic approach to data collection					"There needs to be good practice and procedures in terms of data use."

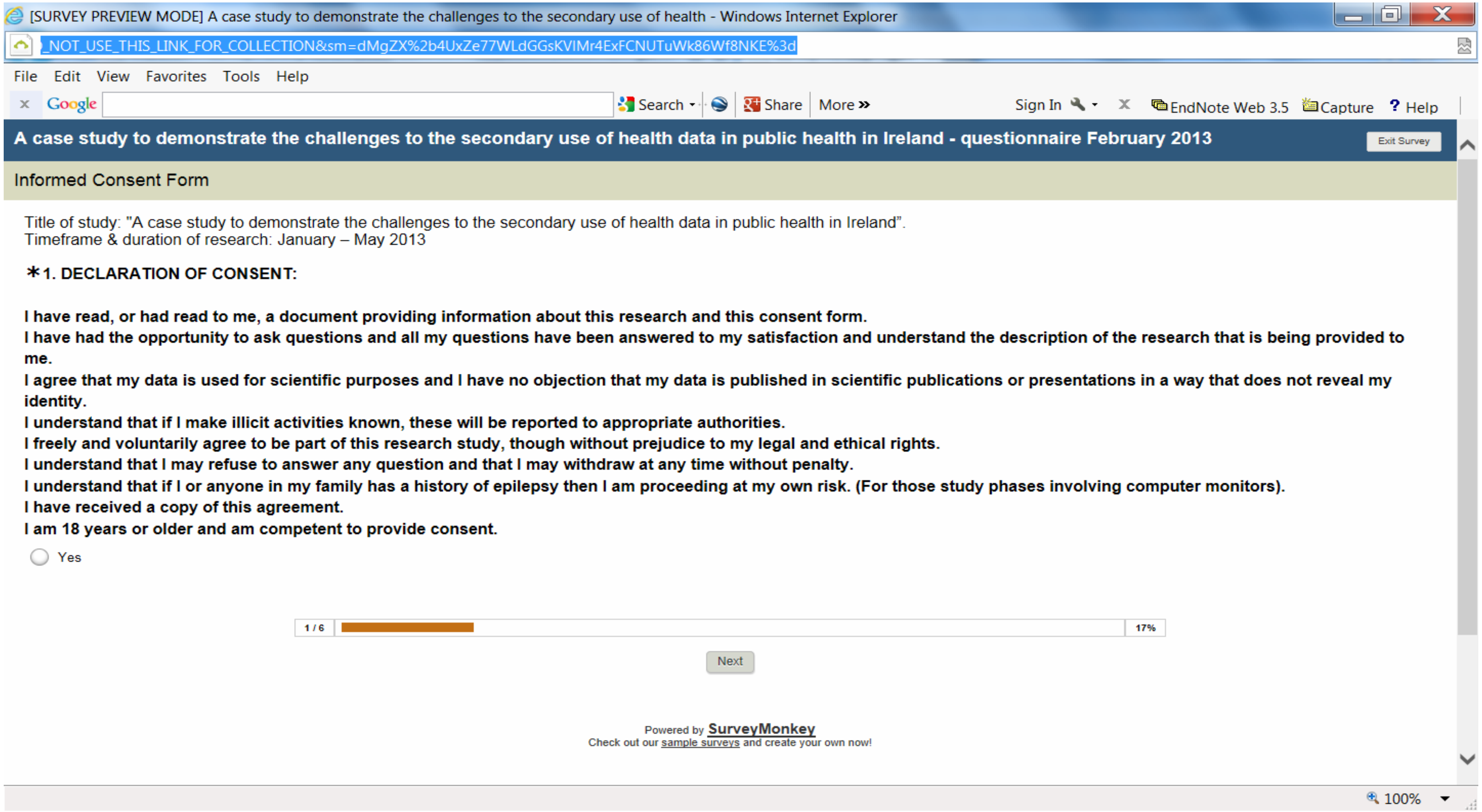
	Interview code	2	5	9	11	17
	Stakeholder group	**Hospital user**	**NCCP**	**HIQA**	**HSE ICT**	**Hospital user**
13e	Better IT sysyetms that can be integrated				From a national perspective there is a view that ICT systems need to be national systems, even in so far as needing only one system for the whole country. There are a number of systems agreed and formally approved and available to hospitals. "Newer systems require a project construction to facilitate agreement on requirements and go to the market place for a framework to meet the requirements that meet the standards of the requirements and thus be interoperable".	
14	**14. Steps to improve national data use**	**14. Steps to improve national data use**	**14. Steps to improve national data use**	**14. Steps to improve national data use**	**14. Steps to improve national data use**	**14. Steps to improve national data use**
14a	Steps to improve data use - reduce reworking of data			There has to be a national view to move forward where different groups must work together on minimum data sets and link with similar developments e.g. HIQA is currently working on a discharge summary and so is the Emergency Care Programme. They have agreed to work together and produce a single discharge summary dataset. Also the work of HIQA on demographic data sets where the Social Services Client Identity Services do a lot of work on this as well. There should be better collaboration to identify those areas where standards and minimum data sets are needed most urgently so that these can be prioritised by HIQA. development of reporting and secondary use cannot wait until all the standards are in place and HIQA are working towards putting the fundamentals in place for these requirements. There is a HIQA advisory group with representatives from the HSE and other sectors, but as the HSE is vast often areas can be		

	Interview code	**2**	**5**	**9**	**11**	**17**
	Stakeholder group	**Hospital user**	**NCCP**	**HIQA**	**HSE ICT**	**Hospital user**
				overlooked.		
14b	Steps to improve data use - incentives/penalties					
14c	Steps to improve data use - data sets, data dictionaries/standards	Agreement of minimum datasets within and between the different clinical programmes.				Minimum data sets and standards
14d	Steps to improve data use - identifiers					Identifiers
14e	Steps to improve data use - use what is collected and collect what is useful					
14f	technology					
14g	Steps to improve data use - legislation			Regulatory framework - the health information bill is very important. The plan is that the Minister will designate some data collections as national collections to which the standards will apply. The question will be to whom they will apply. Patient confidentiality will need to be addressed in the bill to allow the linkage of data from different areas". It is anticipated that the Minister will authorise data matching programmes for sectors such as PCRS, HIPE, NCRI and consent may not be required. The problem will be how will patients "opt out". "Having robust systems that are trusted for secondary use is going to be incredibly important". "This will need a very good robust legal framework that allows this kind of thing (secondary use and data linkage) to go on and that it's done in an open and transparent way and people trust that their information is not being misused or shared inappropriately". Information		

	Interview code	**2**	**5**	**9**	**11**	**17**
	Stakeholder group	**Hospital user**	**NCCP**	**HIQA**	**HSE ICT**	**Hospital user**
				governance issues are a challenge". "There are many reasons why HIQA are strongly opposed to the use of the PPSN here (as the identifier for health care) and believe that there should be a healthcare specific one, one of the issues is there isn't the same kind of confidence amongst the population that their information is going to be kept secure whereas the Scandinavian countries have been using the equivalent of the PPSN for years for everything there are very serious penalties for misuse and data breaches and maybe we have not got to that point and part of it is because we don't have clear accountability as to who is responsible, you know when they find records in bins and bogs and things there's never, or you never hear or get to know, who is responsible, what is the investigation, how did it happen and how can we be confident that it will never happen again" . "People may trust health professionals but not all the other people who have access to their data. electronic systems are good from that perspective because you can control access in a way that is not possible with the paper chart, but then the downside is that if you do manage to break into them then you have access to a vast array of information. For example, if someone got into the PCRS system they would have huge, highly sensitive information.		
14h	Steps to improve data use - pragmatism					

	Interview code	**2**	**5**	**9**	**11**	**17**
	Stakeholder group	**Hospital user**	**NCCP**	**HIQA**	**HSE ICT**	**Hospital user**
14i	Steps to improve data use - leadership/decision making	Leadership/driver required for national standardisation				
15	**15. Comments**	**15. Comments**	**15. Comments**	**15. Comments**	**15. Comments**	**15. Comments**
15	Comments - public/private			Electronic prescribing would facilitate a lot of this data collection and reporting.		
15	Comments - government initiative					

Appendix 9. Hospital pharmacist questionnaire screen shots



A case study to demonstrate the challenges to the secondary use of health data in public health in Ireland - questionnaire February 2013

Informed Consent Form

Title of study: "A case study to demonstrate the challenges to the secondary use of health data in public health in Ireland".
Timeframe & duration of research: January – May 2013

***1. DECLARATION OF CONSENT:**

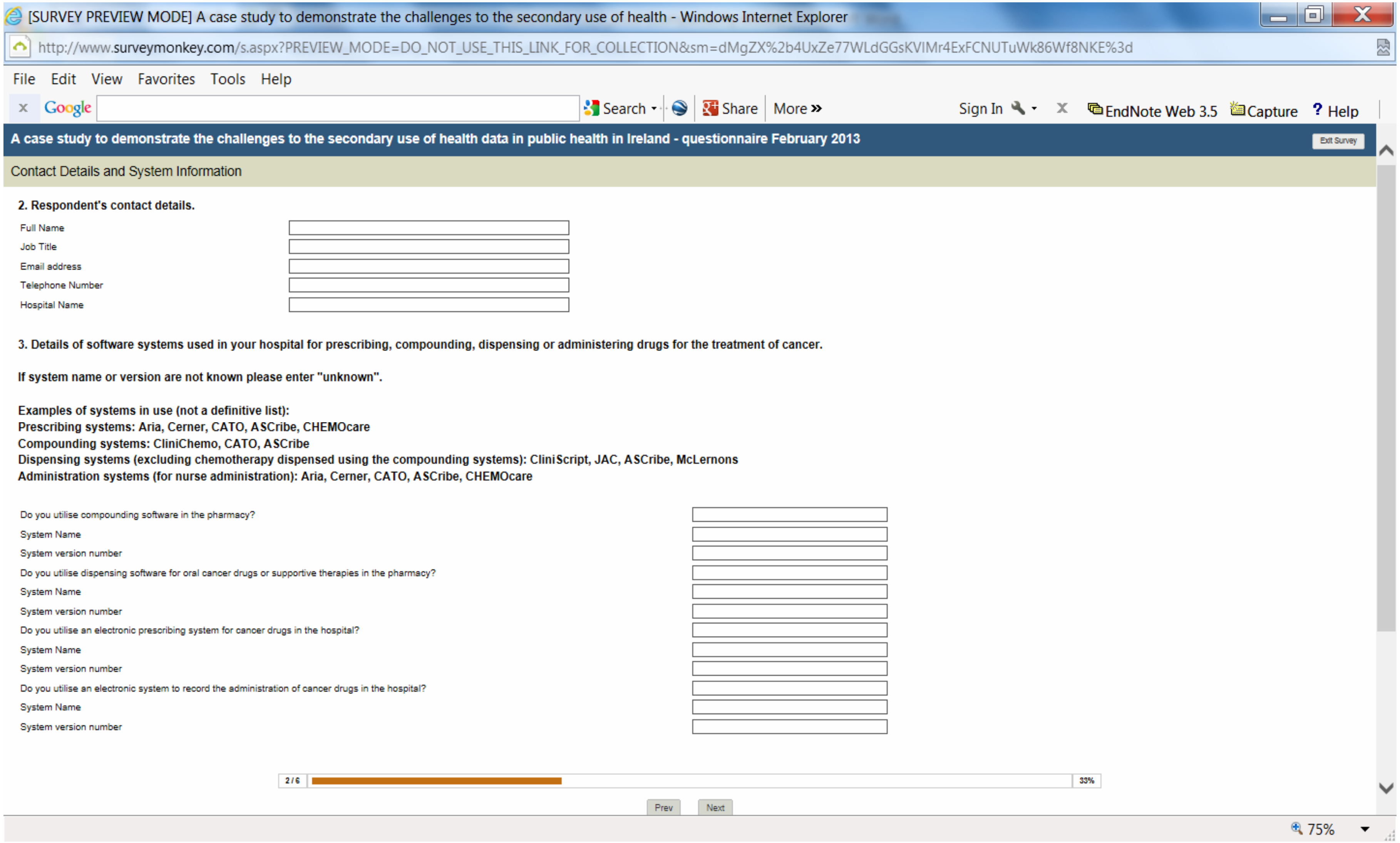
I have read, or had read to me, a document providing information about this research and this consent form.
I have had the opportunity to ask questions and all my questions have been answered to my satisfaction and understand the description of the research that is being provided to me.
I agree that my data is used for scientific purposes and I have no objection that my data is published in scientific publications or presentations in a way that does not reveal my identity.
I understand that if I make illicit activities known, these will be reported to appropriate authorities.
I freely and voluntarily agree to be part of this research study, though without prejudice to my legal and ethical rights.
I understand that I may refuse to answer any question and that I may withdraw at any time without penalty.
I understand that if I or anyone in my family has a history of epilepsy then I am proceeding at my own risk. (For those study phases involving computer monitors).
I have received a copy of this agreement.
I am 18 years or older and am competent to provide consent.

○ Yes

1 / 6 17%

Next

Powered by SurveyMonkey
Check out our sample surveys and create your own now!



A case study to demonstrate the challenges to the secondary use of health data in public health in Ireland - questionnaire February 2013

Contact Details and System Information

2. Respondent's contact details.

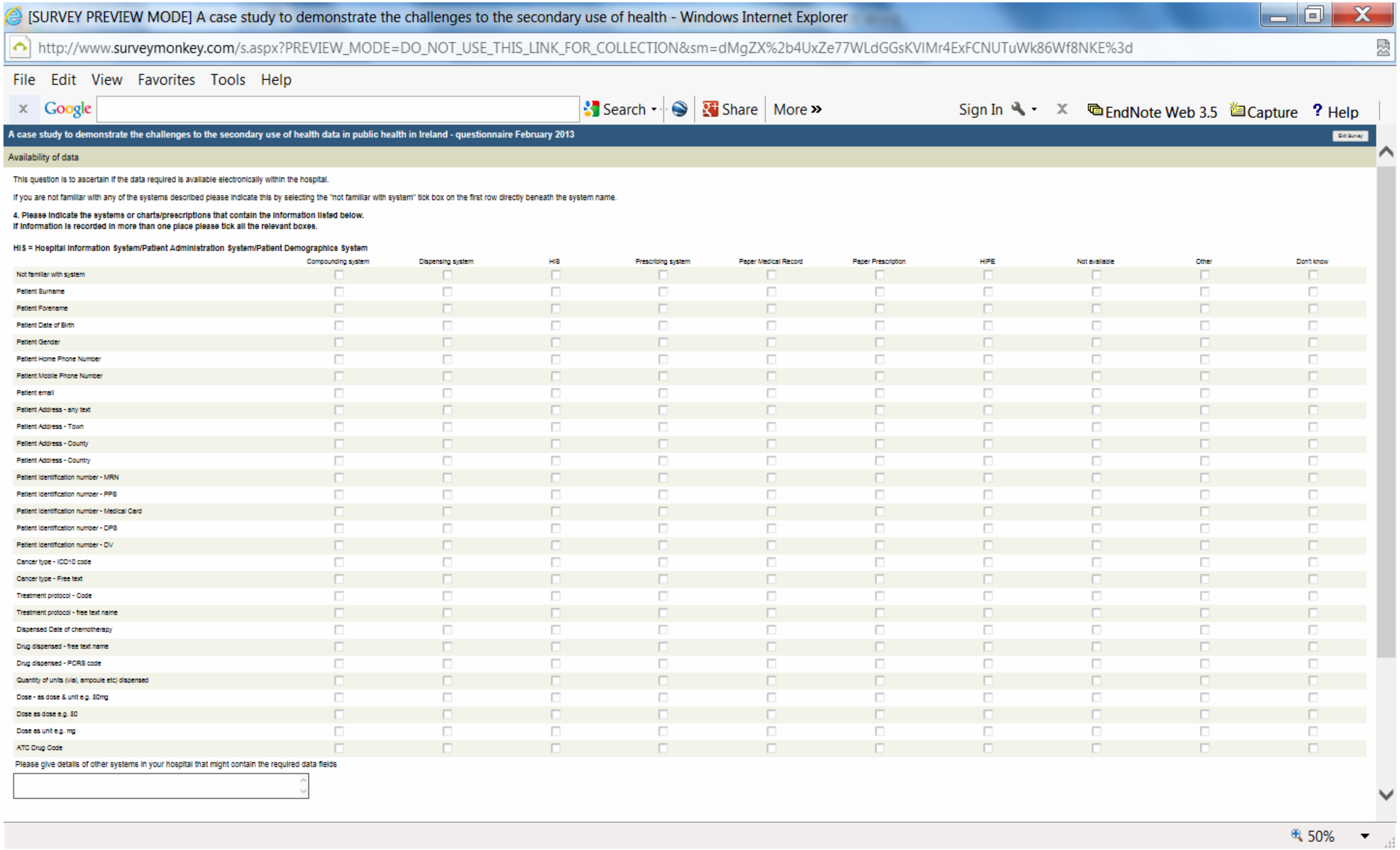
Full Name
Job Title
Email address
Telephone Number
Hospital Name

3. Details of software systems used in your hospital for prescribing, compounding, dispensing or administering drugs for the treatment of cancer.

If system name or version are not known please enter "unknown".

Examples of systems in use (not a definitive list):
Prescribing systems: Aria, Cerner, CATO, ASCribe, CHEMOcare
Compounding systems: CliniChemo, CATO, ASCribe
Dispensing systems (excluding chemotherapy dispensed using the compounding systems): CliniScript, JAC, ASCribe, McLernons
Administration systems (for nurse administration): Aria, Cerner, CATO, ASCribe, CHEMOcare

Do you utilise compounding software in the pharmacy?
System Name
System version number
Do you utilise dispensing software for oral cancer drugs or supportive therapies in the pharmacy?
System Name
System version number
Do you utilise an electronic prescribing system for cancer drugs in the hospital?
System Name
System version number
Do you utilise an electronic system to record the administration of cancer drugs in the hospital?
System Name
System version number



A case study to demonstrate the challenges to the secondary use of health data in public health in Ireland - questionnaire February 2013

Availability of data

This question is to ascertain if the data required is available electronically within the hospital.

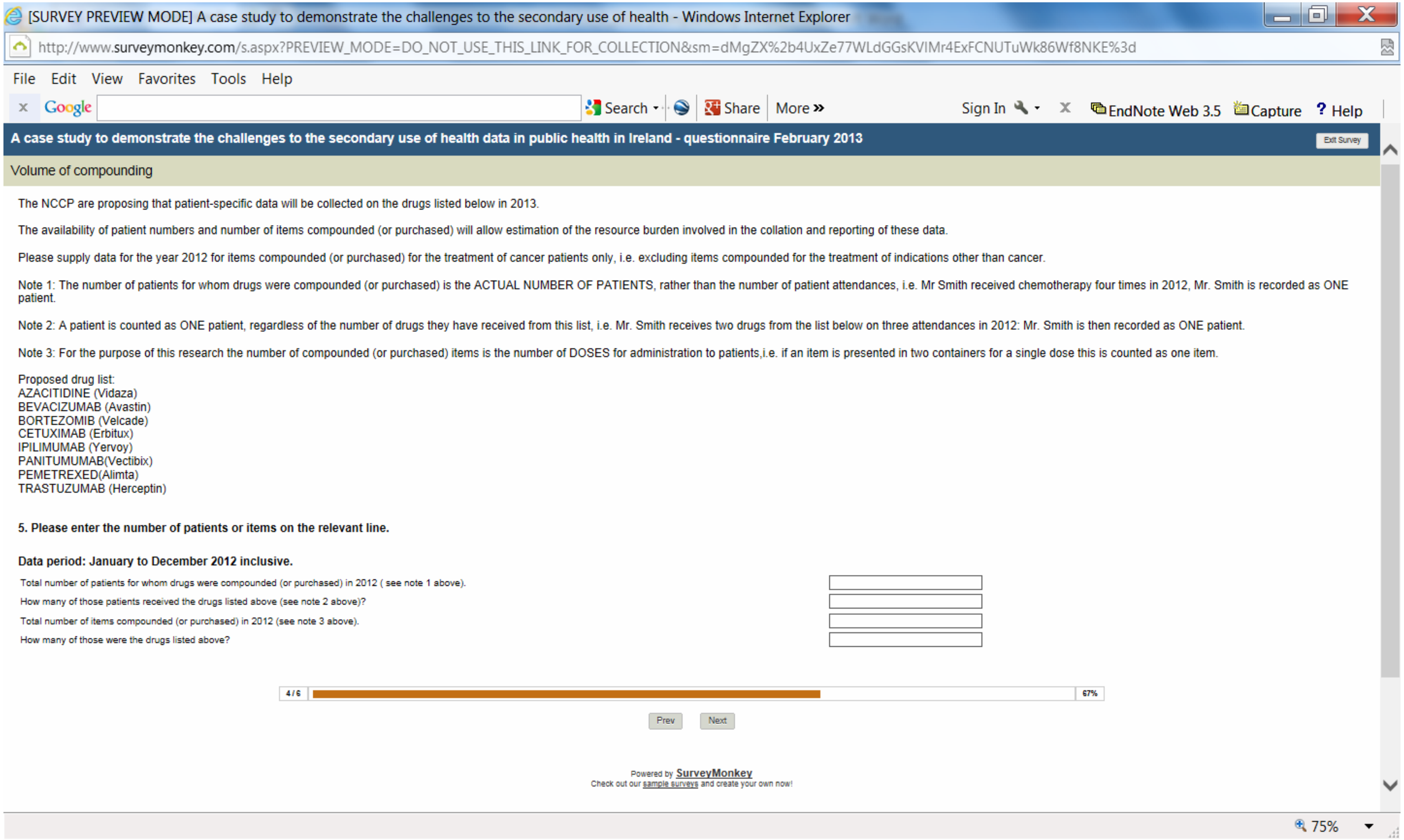
If you are not familiar with any of the systems described please indicate this by selecting the "not familiar with system" tick box on the first row directly beneath the system name.

4. Please indicate the systems or charts/prescriptions that contain the information listed below.
If information is recorded in more than one place please tick all the relevant boxes.

HIS = Hospital Information System/Patient Administration System/Patient Demographics System

	Compounding system	Dispensing system	HIS	Prescribing system	Paper Medical Record	Paper Prescription	HIPE	Not available	Other	Don't know
Not familiar with system	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Patient Surname	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Patient Forename	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Patient Date of Birth	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Patient Gender	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Patient Home Phone Number	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Patient Mobile Phone Number	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Patient email	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Patient Address - any text	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Patient Address - Town	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Patient Address - County	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Patient Address - Country	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Patient identification number - MRN	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Patient identification number - PPS	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Patient identification number - Medical Card	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Patient identification number - DPS	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Patient identification number - DV	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Cancer type - ICD10 code	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Cancer type - Free text	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Treatment protocol - Code	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Treatment protocol - free text name	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Dispensed Date of chemotherapy	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Drug dispensed - free text name	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Drug dispensed - PCRS code	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Quantity of units (vial, ampoule etc) dispensed	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Dose - as dose & unit e.g. 80mg	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Dose as dose e.g. 80	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Dose as unit e.g. mg	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
ATC Drug Code	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Please give details of other systems in your hospital that might contain the required data fields



A case study to demonstrate the challenges to the secondary use of health data in public health in Ireland - questionnaire February 2013

Volume of compounding

The NCCP are proposing that patient-specific data will be collected on the drugs listed below in 2013.

The availability of patient numbers and number of items compounded (or purchased) will allow estimation of the resource burden involved in the collation and reporting of these data.

Please supply data for the year 2012 for items compounded (or purchased) for the treatment of cancer patients only, i.e. excluding items compounded for the treatment of indications other than cancer.

Note 1: The number of patients for whom drugs were compounded (or purchased) is the ACTUAL NUMBER OF PATIENTS, rather than the number of patient attendances, i.e. Mr Smith received chemotherapy four times in 2012, Mr. Smith is recorded as ONE patient.

Note 2: A patient is counted as ONE patient, regardless of the number of drugs they have received from this list, i.e. Mr. Smith receives two drugs from the list below on three attendances in 2012: Mr. Smith is then recorded as ONE patient.

Note 3: For the purpose of this research the number of compounded (or purchased) items is the number of DOSES for administration to patients,i.e. if an item is presented in two containers for a single dose this is counted as one item.

Proposed drug list:
AZACITIDINE (Vidaza)
BEVACIZUMAB (Avastin)
BORTEZOMIB (Velcade)
CETUXIMAB (Erbitux)
IPILIMUMAB (Yervoy)
PANITUMUMAB(Vectibix)
PEMETREXED(Alimta)
TRASTUZUMAB (Herceptin)

5. Please enter the number of patients or items on the relevant line.

Data period: January to December 2012 inclusive.

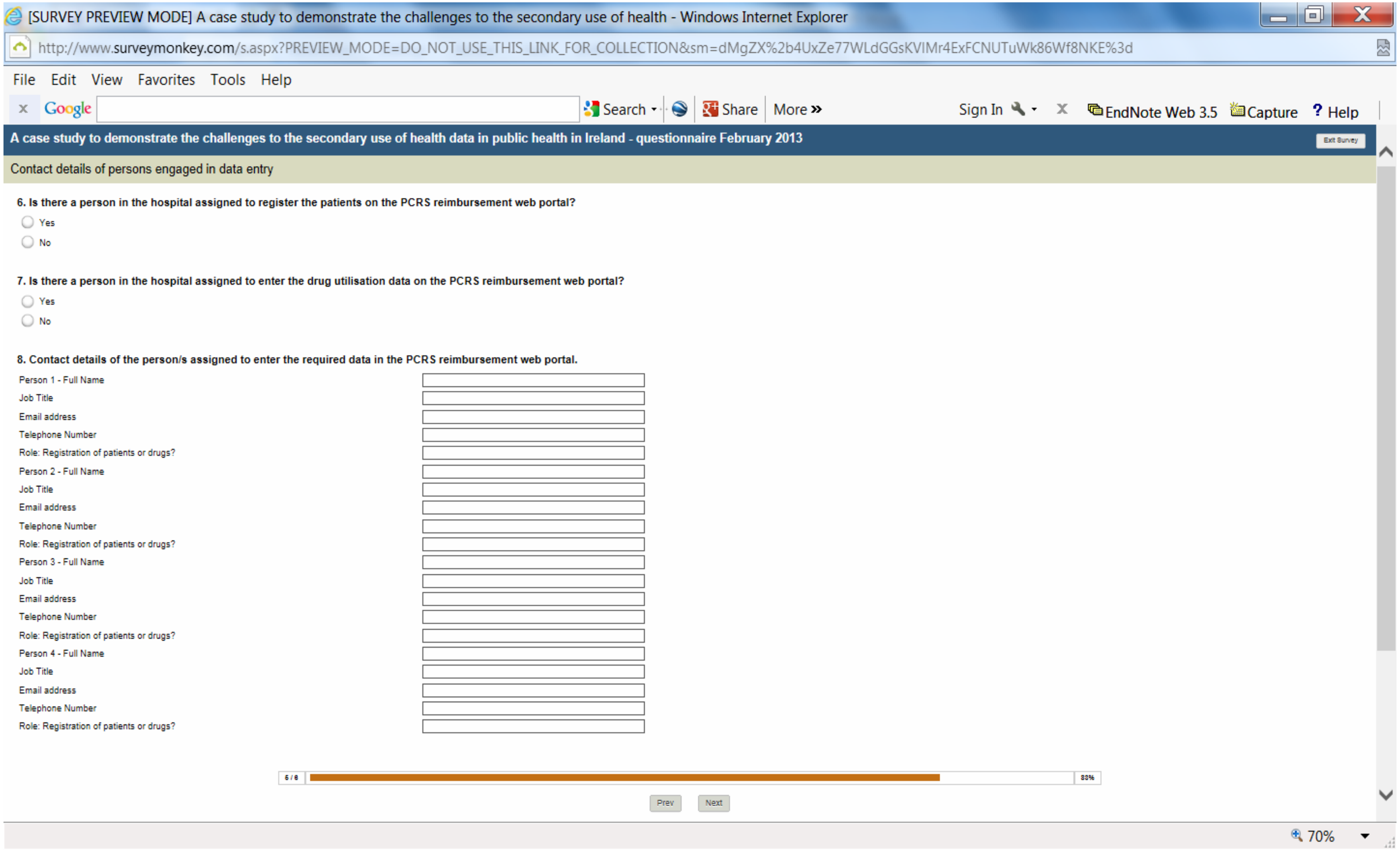
Total number of patients for whom drugs were compounded (or purchased) in 2012 (see note 1 above).

How many of those patients received the drugs listed above (see note 2 above)?

Total number of items compounded (or purchased) in 2012 (see note 3 above).

How many of those were the drugs listed above?

Powered by SurveyMonkey
Check out our sample surveys and create your own now!



A case study to demonstrate the challenges to the secondary use of health data in public health in Ireland - questionnaire February 2013

Contact details of persons engaged in data entry

6. Is there a person in the hospital assigned to register the patients on the PCRS reimbursement web portal?

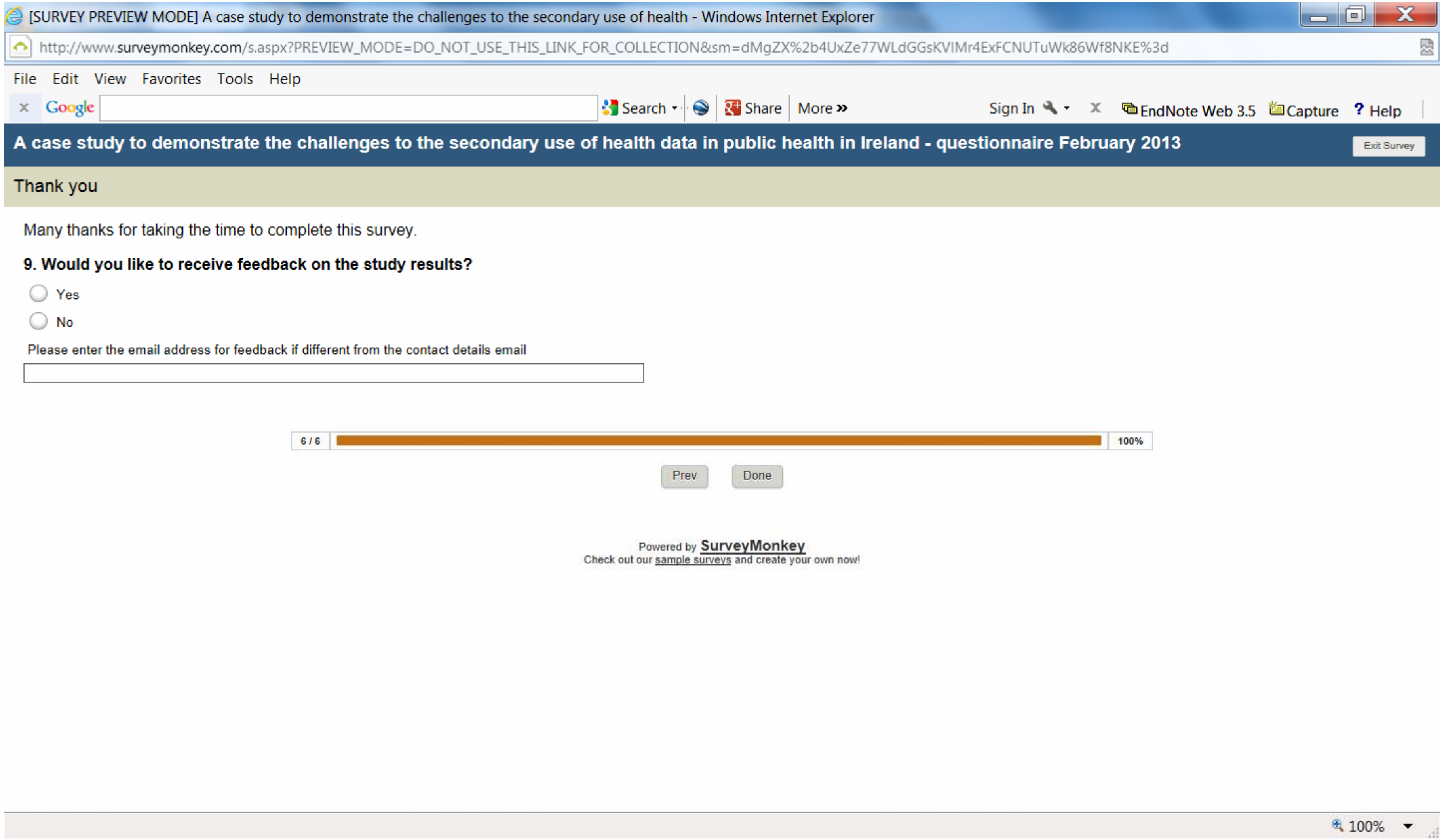
○ Yes
○ No

7. Is there a person in the hospital assigned to enter the drug utilisation data on the PCRS reimbursement web portal?

○ Yes
○ No

8. Contact details of the person/s assigned to enter the required data in the PCRS reimbursement web portal.

Person 1 - Full Name
Job Title
Email address
Telephone Number
Role: Registration of patients or drugs?
Person 2 - Full Name
Job Title
Email address
Telephone Number
Role: Registration of patients or drugs?
Person 3 - Full Name
Job Title
Email address
Telephone Number
Role: Registration of patients or drugs?
Person 4 - Full Name
Job Title
Email address
Telephone Number
Role: Registration of patients or drugs?



A case study to demonstrate the challenges to the secondary use of health data in public health in Ireland - questionnaire February 2013

Thank you

Many thanks for taking the time to complete this survey.

9. Would you like to receive feedback on the study results?

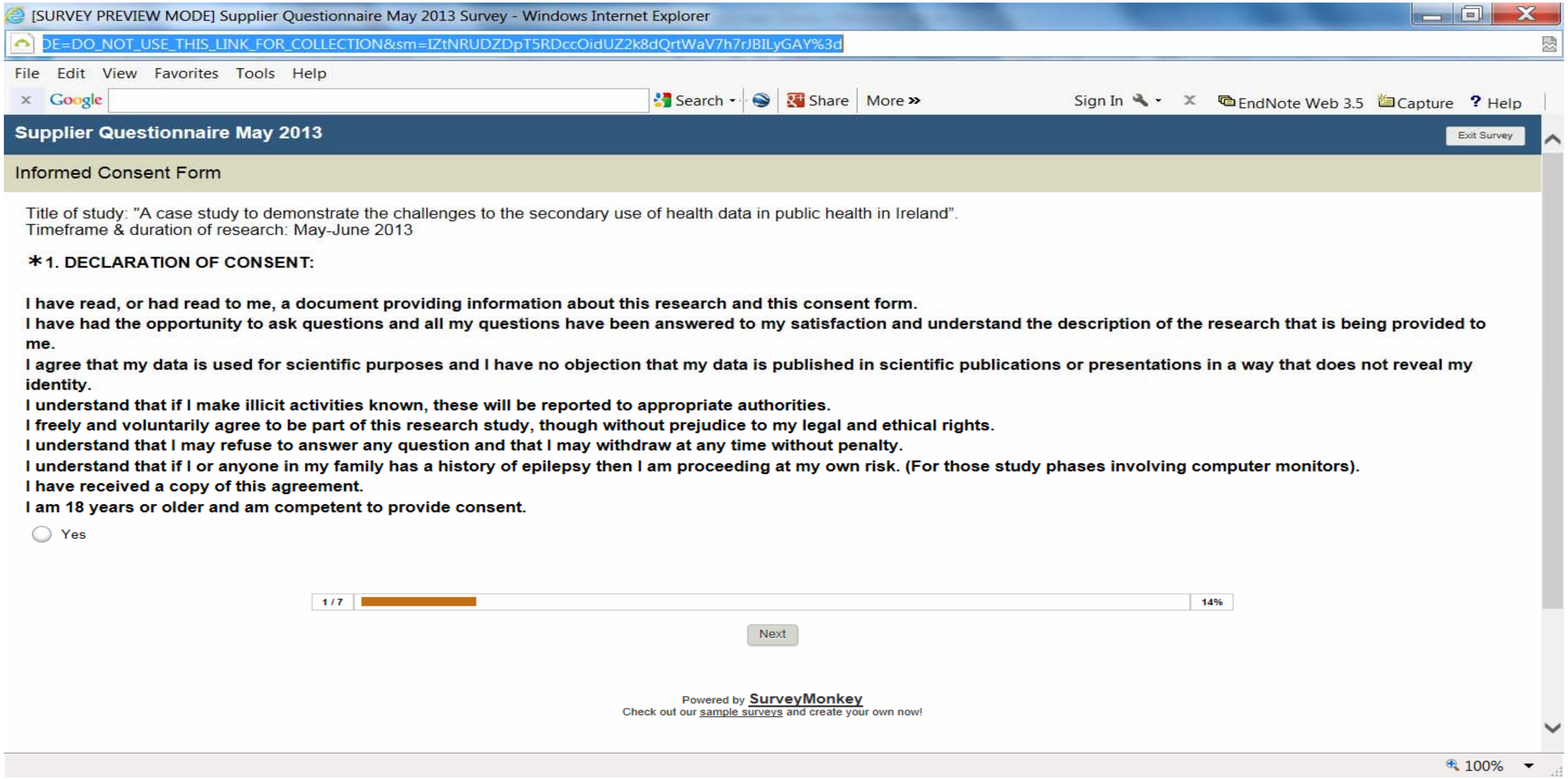
○ Yes

○ No

Please enter the email address for feedback if different from the contact details email

Powered by SurveyMonkey
Check out our sample surveys and create your own now!

# Appendix 10.	Systems vendor questionnaire screen shots



Supplier Questionnaire May 2013

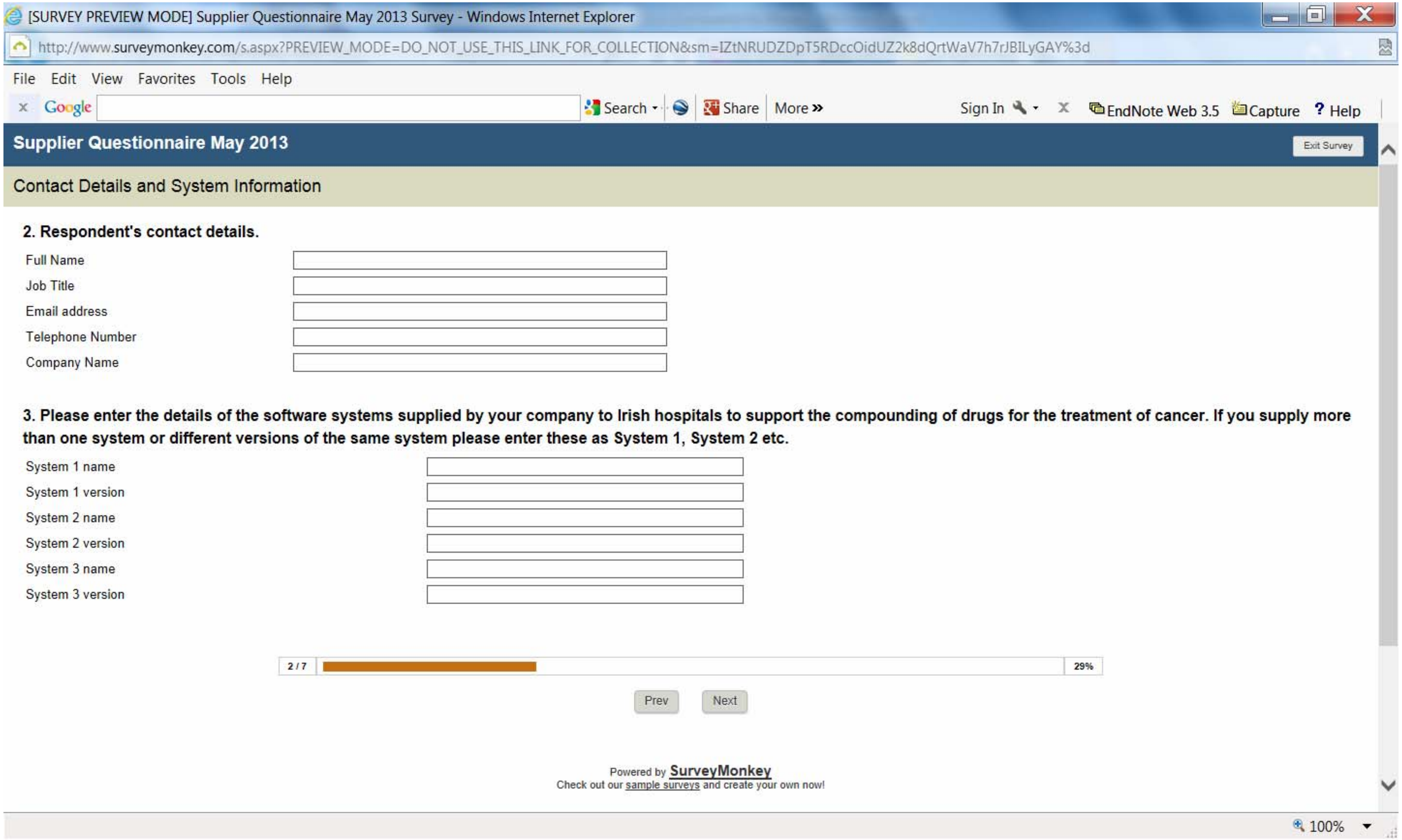
Informed Consent Form

Title of study: "A case study to demonstrate the challenges to the secondary use of health data in public health in Ireland".
Timeframe & duration of research: May-June 2013

***1. DECLARATION OF CONSENT:**

I have read, or had read to me, a document providing information about this research and this consent form.
I have had the opportunity to ask questions and all my questions have been answered to my satisfaction and understand the description of the research that is being provided to me.
I agree that my data is used for scientific purposes and I have no objection that my data is published in scientific publications or presentations in a way that does not reveal my identity.
I understand that if I make illicit activities known, these will be reported to appropriate authorities.
I freely and voluntarily agree to be part of this research study, though without prejudice to my legal and ethical rights.
I understand that I may refuse to answer any question and that I may withdraw at any time without penalty.
I understand that if I or anyone in my family has a history of epilepsy then I am proceeding at my own risk. (For those study phases involving computer monitors).
I have received a copy of this agreement.
I am 18 years or older and am competent to provide consent.

○ Yes



Supplier Questionnaire May 2013

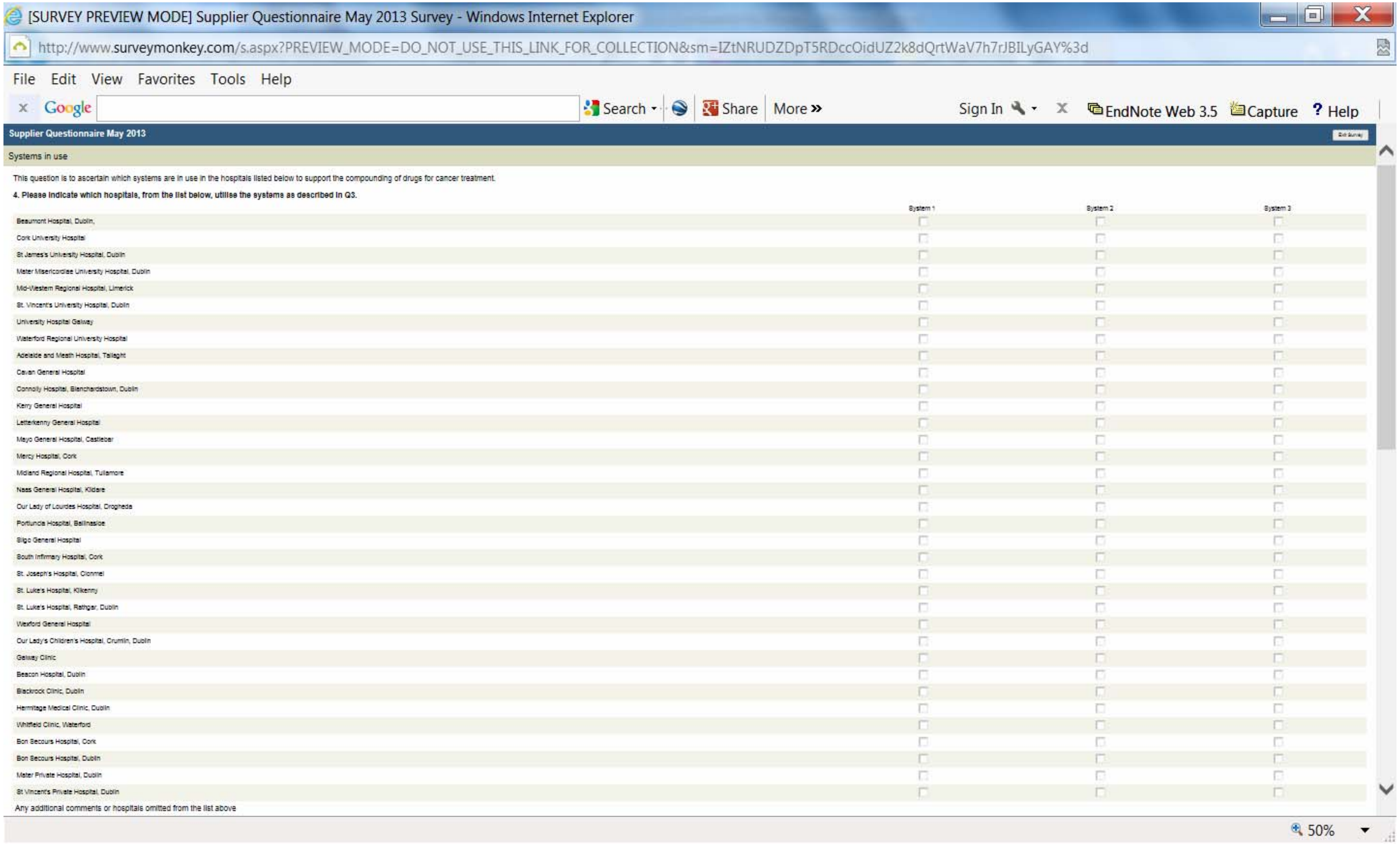
Contact Details and System Information

2. Respondent's contact details.

Full Name
Job Title
Email address
Telephone Number
Company Name

3. Please enter the details of the software systems supplied by your company to Irish hospitals to support the compounding of drugs for the treatment of cancer. If you supply more than one system or different versions of the same system please enter these as System 1, System 2 etc.

System 1 name
System 1 version
System 2 name
System 2 version
System 3 name
System 3 version



Supplier Questionnaire May 2013

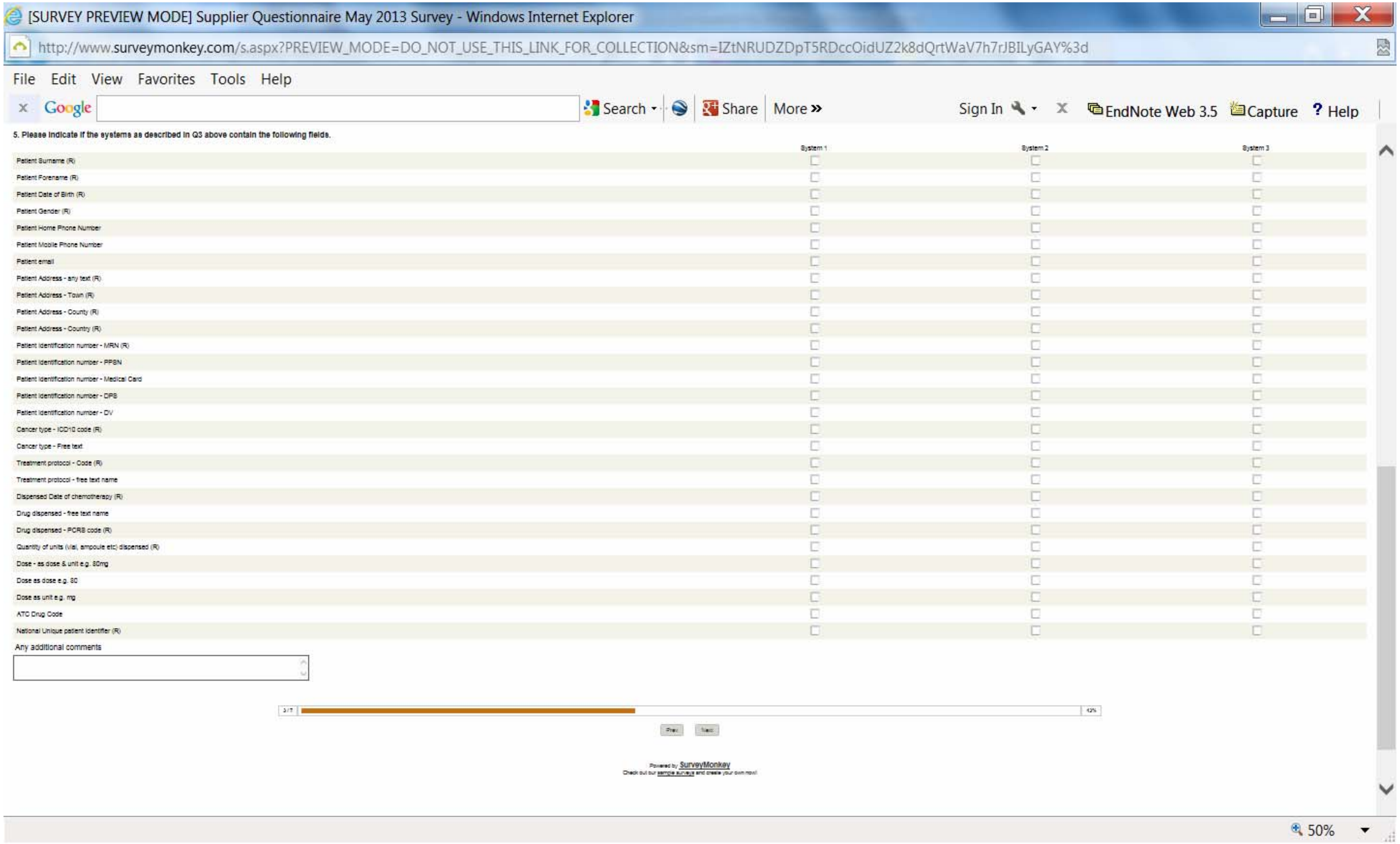
Systems in use

This question is to ascertain which systems are in use in the hospitals listed below to support the compounding of drugs for cancer treatment.

4. Please indicate which hospitals, from the list below, utilise the systems as described in Q3.

	System 1	System 2	System 3
Beaumont Hospital, Dublin,	☐	☐	☐
Cork University Hospital	☐	☐	☐
St James's University Hospital, Dublin	☐	☐	☐
Mater Misericordiae University Hospital, Dublin	☐	☐	☐
Mid-Western Regional Hospital, Limerick	☐	☐	☐
St. Vincent's University Hospital, Dublin	☐	☐	☐
University Hospital Galway	☐	☐	☐
Waterford Regional University Hospital	☐	☐	☐
Adelaide and Meath Hospital, Tallaght	☐	☐	☐
Cavan General Hospital	☐	☐	☐
Connolly Hospital, Blanchardstown, Dublin	☐	☐	☐
Kerry General Hospital	☐	☐	☐
Letterkenny General Hospital	☐	☐	☐
Mayo General Hospital, Castlebar	☐	☐	☐
Mercy Hospital, Cork	☐	☐	☐
Midland Regional Hospital, Tullamore	☐	☐	☐
Naas General Hospital, Kildare	☐	☐	☐
Our Lady of Lourdes Hospital, Drogheda	☐	☐	☐
Portiuncla Hospital, Ballinasloe	☐	☐	☐
Sligo General Hospital	☐	☐	☐
South Infirmary Hospital, Cork	☐	☐	☐
St. Joseph's Hospital, Clonmel	☐	☐	☐
St. Luke's Hospital, Kilkenny	☐	☐	☐
St. Luke's Hospital, Rathgar, Dublin	☐	☐	☐
Wexford General Hospital	☐	☐	☐
Our Lady's Children's Hospital, Crumlin, Dublin	☐	☐	☐
Galway Clinic	☐	☐	☐
Beacon Hospital, Dublin	☐	☐	☐
Blackrock Clinic, Dublin	☐	☐	☐
Hermitage Medical Clinic, Dublin	☐	☐	☐
Whitfield Clinic, Waterford	☐	☐	☐
Bon Secours Hospital, Cork	☐	☐	☐
Bon Secours Hospital, Dublin	☐	☐	☐
Mater Private Hospital, Dublin	☐	☐	☐
St Vincent's Private Hospital, Dublin	☐	☐	☐

Any additional comments or hospitals omitted from the list above



[SURVEY PREVIEW MODE] Supplier Questionnaire May 2013 Survey - Windows Internet Explorer

http://www.surveymonkey.com/s.aspx?PREVIEW_MODE=DO_NOT_USE_THIS_LINK_FOR_COLLECTION&sm=IZtNRUDZDpT5RDccOidUZ2k8dQrtWaV7h7rJBILyGAY%3d

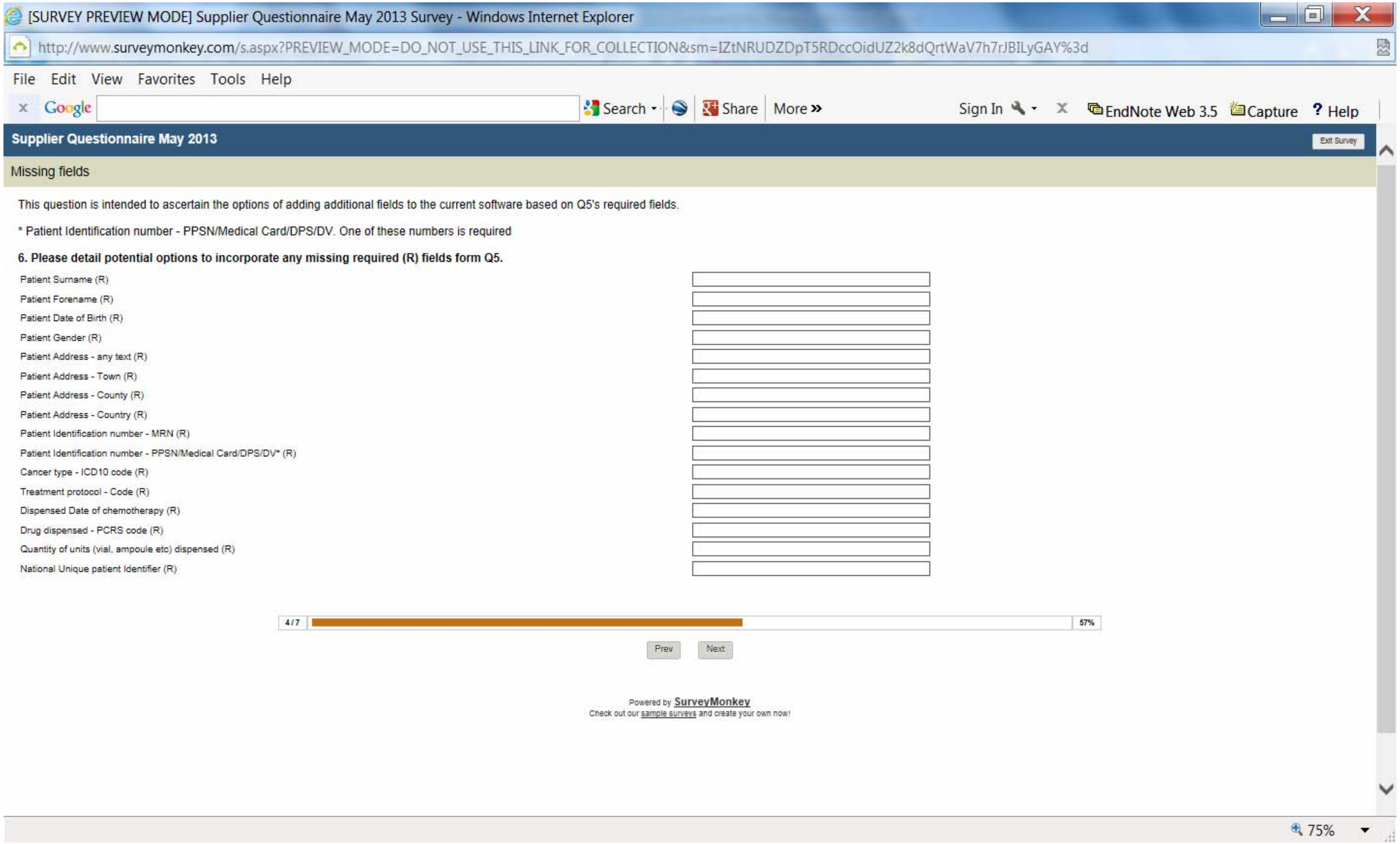
5. Please indicate if the systems as described in Q3 above contain the following fields.

	System 1	System 2	System 3
Patient Surname (R)	☐	☐	☐
Patient Forename (R)	☐	☐	☐
Patient Date of Birth (R)	☐	☐	☐
Patient Gender (R)	☐	☐	☐
Patient Home Phone Number	☐	☐	☐
Patient Mobile Phone Number	☐	☐	☐
Patient email	☐	☐	☐
Patient Address - any text (R)	☐	☐	☐
Patient Address - Town (R)	☐	☐	☐
Patient Address - County (R)	☐	☐	☐
Patient Address - Country (R)	☐	☐	☐
Patient identification number - MRN (R)	☐	☐	☐
Patient identification number - PPSN	☐	☐	☐
Patient identification number - Medical Card	☐	☐	☐
Patient identification number - DPS	☐	☐	☐
Patient identification number - DV	☐	☐	☐
Cancer type - ICD10 code (R)	☐	☐	☐
Cancer type - Free text	☐	☐	☐
Treatment protocol - Code (R)	☐	☐	☐
Treatment protocol - free text name	☐	☐	☐
Dispensed Date of chemotherapy (R)	☐	☐	☐
Drug dispensed - free text name	☐	☐	☐
Drug dispensed - PCRS code (R)	☐	☐	☐
Quantity of units (vial, ampoule etc) dispensed (R)	☐	☐	☐
Dose - as dose & unit e.g. 80mg	☐	☐	☐
Dose as dose e.g. 80	☐	☐	☐
Dose as unit e.g. mg	☐	☐	☐
ATC Drug Code	☐	☐	☐
National Unique patient identifier (R)	☐	☐	☐

Any additional comments

3 / 7 42%

Powered by SurveyMonkey
Check out our sample surveys and create your own now!



Supplier Questionnaire May 2013

Missing fields

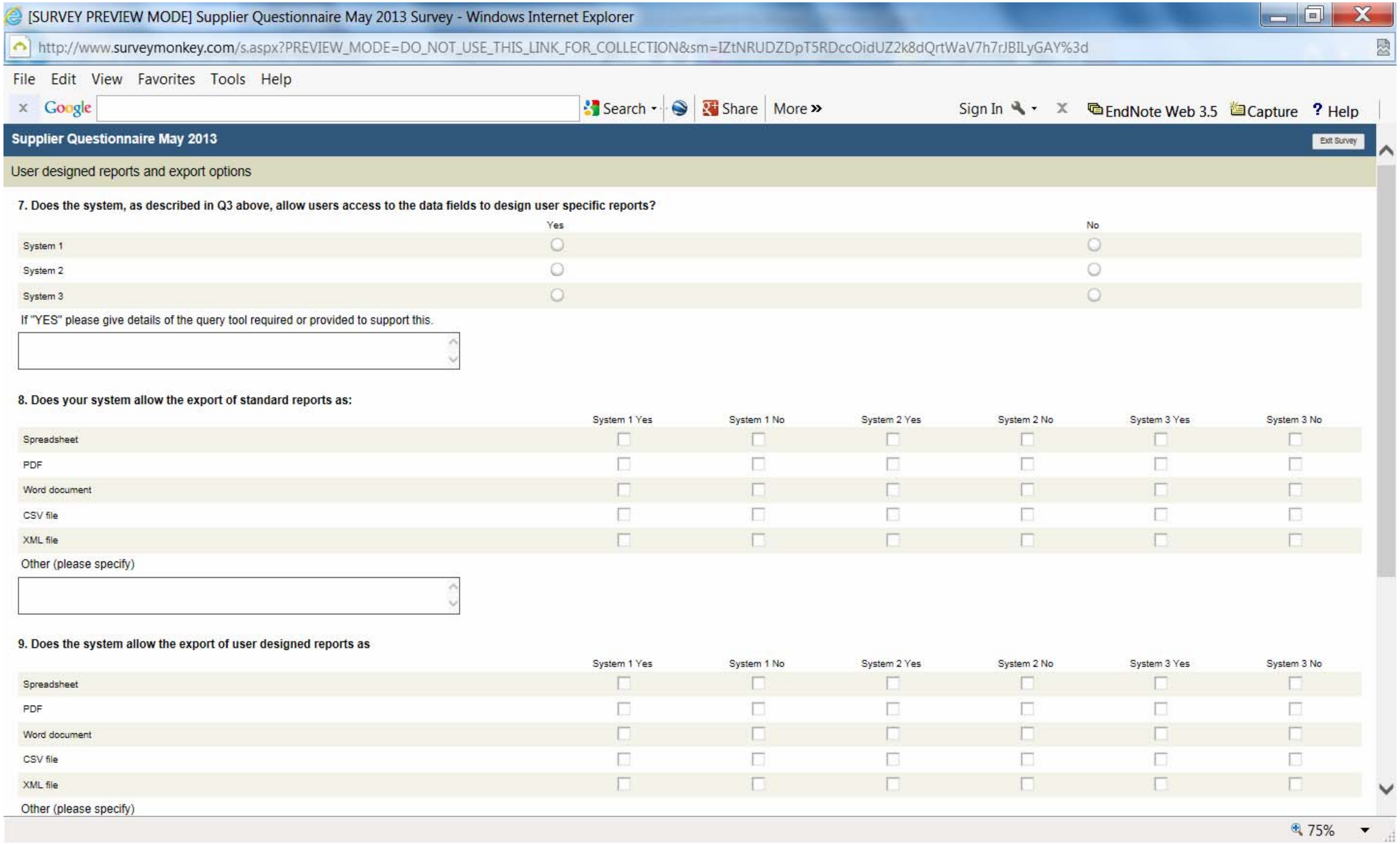
This question is intended to ascertain the options of adding additional fields to the current software based on Q5's required fields.

* Patient Identification number - PPSN/Medical Card/DPS/DV. One of these numbers is required

6. Please detail potential options to incorporate any missing required (R) fields form Q5.

Patient Surname (R)
Patient Forename (R)
Patient Date of Birth (R)
Patient Gender (R)
Patient Address - any text (R)
Patient Address - Town (R)
Patient Address - County (R)
Patient Address - Country (R)
Patient Identification number - MRN (R)
Patient Identification number - PPSN/Medical Card/DPS/DV* (R)
Cancer type - ICD10 code (R)
Treatment protocol - Code (R)
Dispensed Date of chemotherapy (R)
Drug dispensed - PCRS code (R)
Quantity of units (vial, ampoule etc) dispensed (R)
National Unique patient Identifier (R)

Powered by SurveyMonkey
Check out our sample surveys and create your own now!



Supplier Questionnaire May 2013

User designed reports and export options

7. Does the system, as described in Q3 above, allow users access to the data fields to design user specific reports?

	Yes	No
System 1	○	○
System 2	○	○
System 3	○	○

If "YES" please give details of the query tool required or provided to support this.

8. Does your system allow the export of standard reports as:

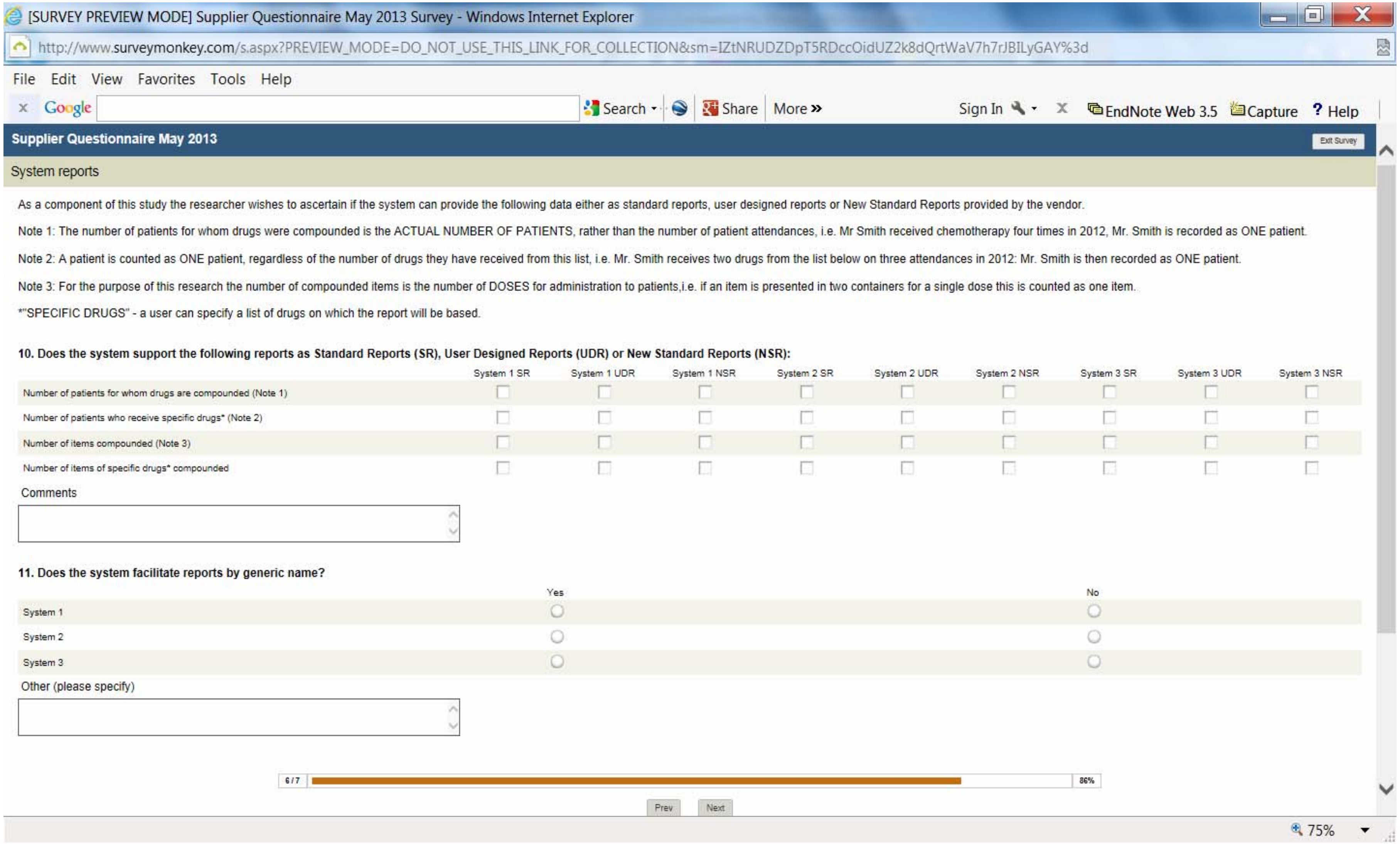
	System 1 Yes	System 1 No	System 2 Yes	System 2 No	System 3 Yes	System 3 No
Spreadsheet	☐	☐	☐	☐	☐	☐
PDF	☐	☐	☐	☐	☐	☐
Word document	☐	☐	☐	☐	☐	☐
CSV file	☐	☐	☐	☐	☐	☐
XML file	☐	☐	☐	☐	☐	☐

Other (please specify)

9. Does the system allow the export of user designed reports as

	System 1 Yes	System 1 No	System 2 Yes	System 2 No	System 3 Yes	System 3 No
Spreadsheet	☐	☐	☐	☐	☐	☐
PDF	☐	☐	☐	☐	☐	☐
Word document	☐	☐	☐	☐	☐	☐
CSV file	☐	☐	☐	☐	☐	☐
XML file	☐	☐	☐	☐	☐	☐

Other (please specify)



Supplier Questionnaire May 2013

System reports

As a component of this study the researcher wishes to ascertain if the system can provide the following data either as standard reports, user designed reports or New Standard Reports provided by the vendor.

Note 1: The number of patients for whom drugs were compounded is the ACTUAL NUMBER OF PATIENTS, rather than the number of patient attendances, i.e. Mr Smith received chemotherapy four times in 2012, Mr. Smith is recorded as ONE patient.

Note 2: A patient is counted as ONE patient, regardless of the number of drugs they have received from this list, i.e. Mr. Smith receives two drugs from the list below on three attendances in 2012: Mr. Smith is then recorded as ONE patient.

Note 3: For the purpose of this research the number of compounded items is the number of DOSES for administration to patients,i.e. if an item is presented in two containers for a single dose this is counted as one item.

*"SPECIFIC DRUGS" - a user can specify a list of drugs on which the report will be based.

10. Does the system support the following reports as Standard Reports (SR), User Designed Reports (UDR) or New Standard Reports (NSR):

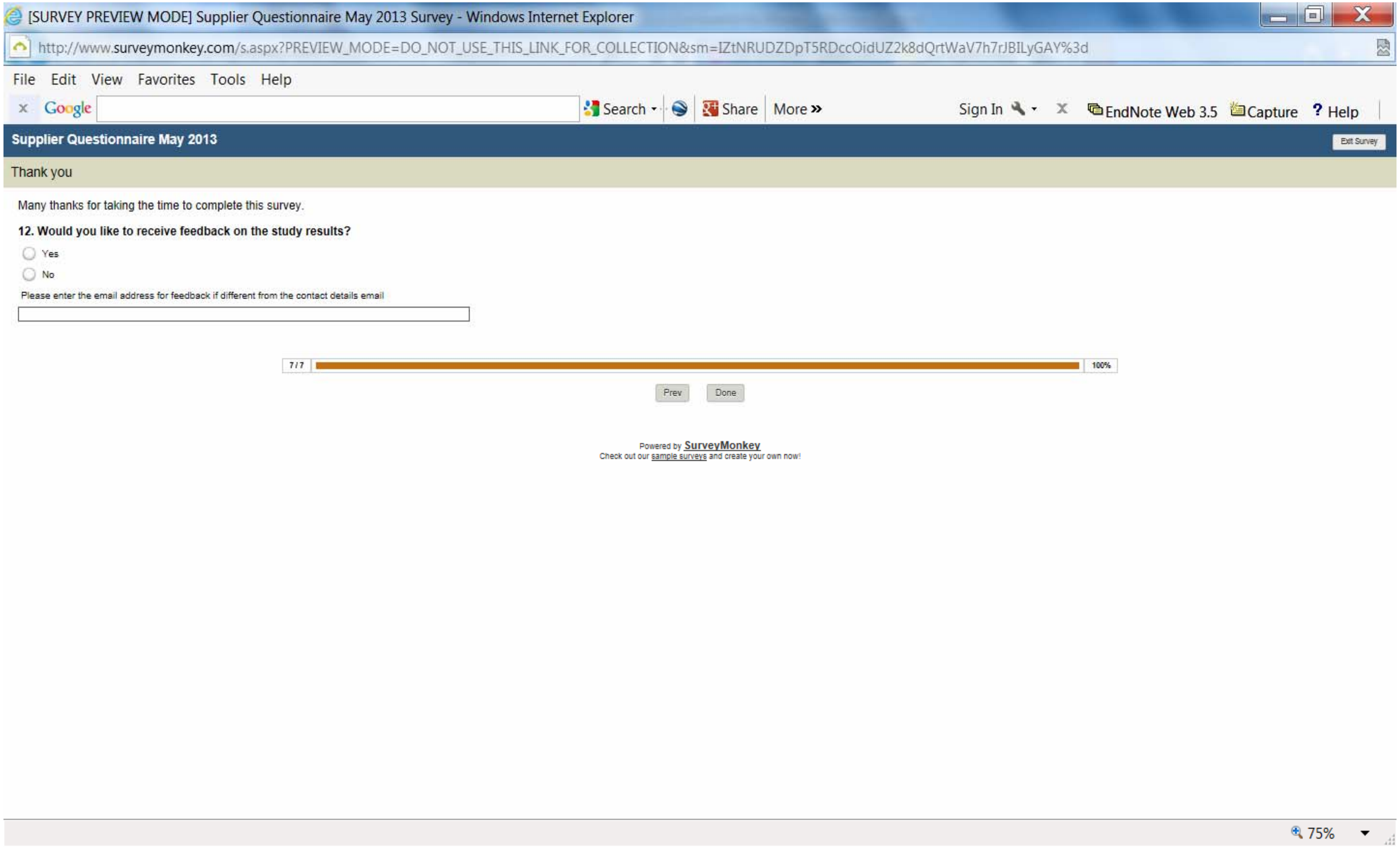
	System 1 SR	System 1 UDR	System 1 NSR	System 2 SR	System 2 UDR	System 2 NSR	System 3 SR	System 3 UDR	System 3 NSR
Number of patients for whom drugs are compounded (Note 1)	☐	☐	☐	☐	☐	☐	☐	☐	☐
Number of patients who receive specific drugs* (Note 2)	☐	☐	☐	☐	☐	☐	☐	☐	☐
Number of items compounded (Note 3)	☐	☐	☐	☐	☐	☐	☐	☐	☐
Number of items of specific drugs* compounded	☐	☐	☐	☐	☐	☐	☐	☐	☐

Comments

11. Does the system facilitate reports by generic name?

	Yes	No
System 1	○	○
System 2	○	○
System 3	○	○

Other (please specify)



Supplier Questionnaire May 2013

Thank you

Many thanks for taking the time to complete this survey.

12. Would you like to receive feedback on the study results?

○ Yes

○ No

Please enter the email address for feedback if different from the contact details email

Powered by SurveyMonkey
Check out our sample surveys and create your own now!

Appendix 11. Location of data in hospital systems

Results downloaded from SurveyMonkey®.

Please indicate the systems or charts/prescriptions that contain the information listed below. If information is recorded in more than one place please tick all the relevant boxes. HI
System/Patient Administration System/Patient Demographics System

Answer Options	Compounding system	Dispensing system	HIS	Prescribing system	Paper Medical	Paper Prescription	HIPE	Not available	Other
Not familiar with system	2	0	8	6	2	0	14	1	0
Patient Surname	17	21	14	5	20	21	6	0	0
Patient Forename	16	20	13	5	20	21	6	0	0
Patient Date of Birth	15	16	13	5	19	21	5	0	0
Patient Gender	15	13	13	5	18	12	3	0	0
Patient Home Phone Number	1	0	12	3	12	0	0	2	0
Patient Mobile Phone Number	1	0	11	3	10	0	0	2	0
Patient email	1	0	1	3	3	0	0	7	0
Patient Address - any text	10	10	13	4	18	11	3	0	1
Patient Address - Town	10	10	13	4	17	10	3	0	1
Patient Address - County	8	9	13	4	17	10	3	0	1
Patient Address - Country	6	4	7	3	11	5	1	0	1
Patient Identification number - MRN	16	17	12	5	19	19	4	0	0
Patient Identification number - PPS	1	0	5	2	4	1	0	3	0
Patient Identification number - Medical Card	1	0	7	2	4	2	0	2	0
Patient Identification number - DPS	0	0	3	1	2	1	0	5	0
Patient Identification number - DV	0	0	0	0	0	0	0	3	0
Cancer type - ICD10 code	1	0	1	4	3	0	3	3	0
Cancer type - Free text	1	3	0	2	16	10	2	0	0
Treatment protocol - Code	9	1	0	3	2	2	1	2	0
Treatment protocol - free text name	9	3	0	2	13	14	0	0	1
Dispensed Date of chemotherapy	17	16	0	3	12	11	0	0	0
Drug dispensed - free text name	16	18	0	3	11	14	1	0	0
Drug dispensed - PCRS code	1	1	0	0	0	0	0	6	1
Quantity of units (vial, ampoule etc) dispensed	12	14	0	1	1	2	0	1	1
Dose - as dose & unit e.g. 80mg	16	14	0	5	9	15	1	0	0
Dose as dose e.g. 80	6	5	0	4	2	6	0	0	1
Dose as unit e.g. mg	6	5	0	4	2	6	0	0	0
ATC Drug Code	2	3	0	0	0	0	0	4	2

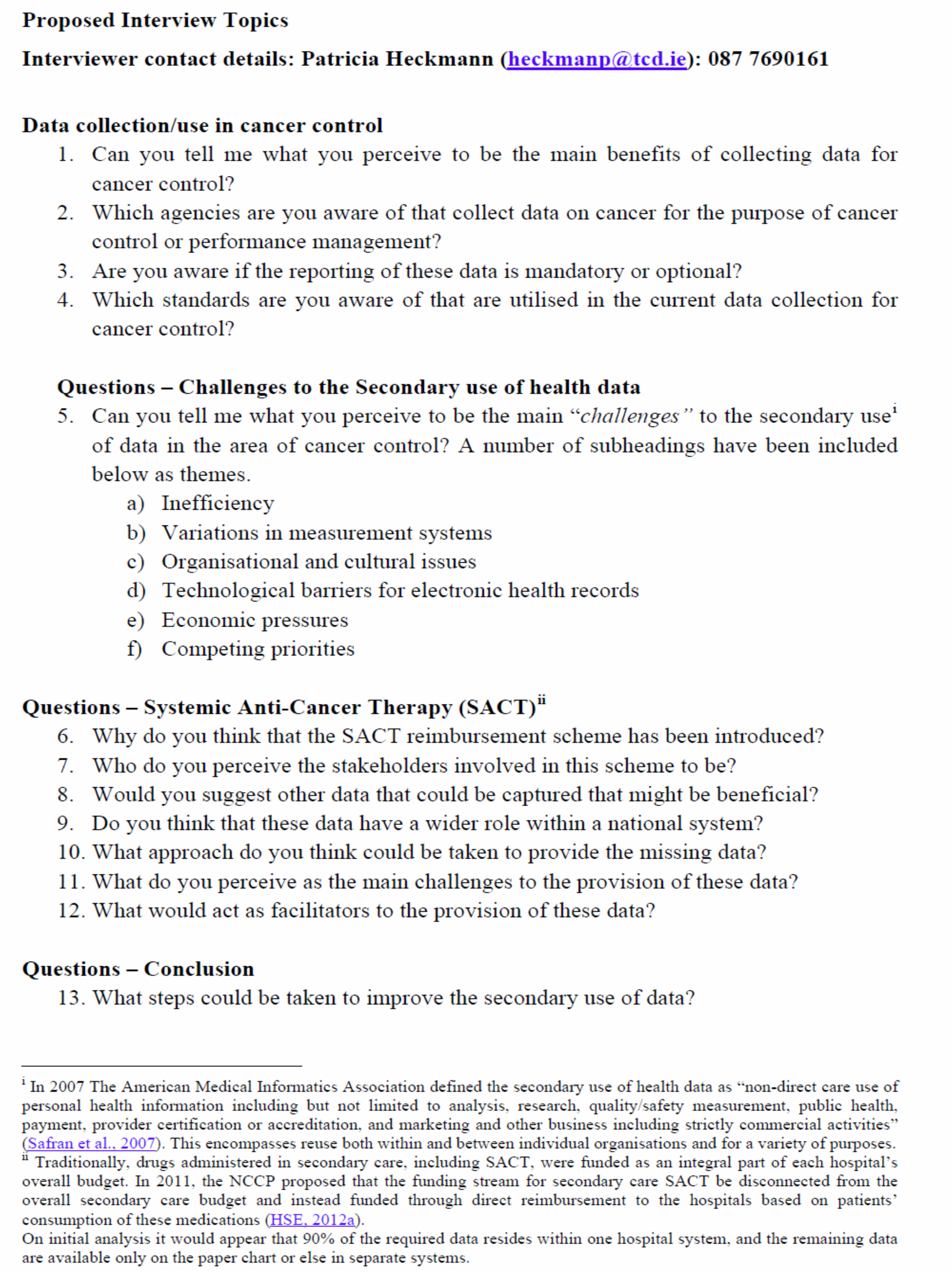
Appendix 12. Hospital and NCCP data responses

Cancer Center	System Code	Hospital Code	Number of patients		Number of patients on specified drugs		Number of doses		Number of doses of specified drugs		New patients
			NCCP	Survey	NCCP	Survey	NCCP	Survey	NCCP	Survey	NCCP
N	Dispensing System	16	Data	Data	Data	Data	Data	Data	Data	Data	Data
N	System A	19	Data	No reply	Data	Data	Data	No reply	Data	No reply	Data
N	System A	1	Data	Data	Data	Data	Data	Data	Data	Data	Data
N	System A	22	Data	No reply	Data	No reply	Data	No reply	Data	No reply	Data
N	System A	17	Data	No reply	Data	No reply	Data	No reply	Data	No reply	Data
N	System A	20	Data	Data	Data	Data	Data	Data	Data	Data	Data
N	System A	7	Data	Data	Data	Data	Data	Data	Data	Data	Data
N	System A	25	Data	No reply	Data	No reply	Data	No reply	Data	No reply	Data
N	System A	9	No reply	Data	No reply	Data	No reply	Data	No reply	Data	No reply
N	System A	18	No reply	No reply	Data	No reply	No reply	No reply	No reply	No reply	No reply
N	System A	15	Data	Data	Data	Data	Data	Data	Data	Data	Data
N	System A	11	Data	Data	Data	Data	Data	Data	Data	Data	Data
N	Dispensing System	23	Data	Data	Data	Data	Data	Data	Data	Data	Data
N	System A	21	Data	No reply	Data	No reply	Data	No reply	Data	No reply	Data
N	System A	13	No reply	No reply	No reply	No reply	No reply	No reply	No reply	No reply	No reply
N	System A	24	Data	No reply	Data	No reply	Data	No reply	Data	No reply	Data
N	System A	6	No reply	Data	No reply	Data	No reply	Data	No reply	Data	No reply
N	System A	2	No reply	Data	No reply	Data	No reply	Data	No reply	Data	No reply
Y	System D	12	Data	Data	Data	Data	Data	Data	Data	Data	Data
Y	System D	5	Data	Data	Data	Data	Data	Data	Data	Data	Data
Y	System B	10	Data	Data	Data	Data	Data	Data	Data	Data	Data
Y	System C	8	Data	Data	Data	Data	Data	Data	Data	Data	Data
Y	System A	14	Data	Data	Data	Data	Data	Data	Data	Data	Data
Y	System A	4	Data	Data	Data	Data	Data	Data	Data	Data	Data
Y	System A	26	Data	No reply	Data	No reply	Data	No reply	Data	No reply	Data
Y	System A	3	Data	No reply	Data	No reply	Data	No reply	No reply	No reply	Data

Appendix 13. Hospitals included in the research

Hospital	**Category**
Beaumont Hospital, Dublin	Specialist Cancer Centre
Cork University Hospital	Specialist Cancer Centre
St James's University Hospital, Dublin	Specialist Cancer Centre
Mater Misericordiae University Hospital, Dublin	Specialist Cancer Centre
Mid-Western Regional Hospital, Limerick	Specialist Cancer Centre
St. Vincent's University Hospital, Dublin	Specialist Cancer Centre
University Hospital Galway	Specialist Cancer Centre
Waterford Regional University Hospital	Specialist Cancer Centre
Adelaide and Meath Hospital, Tallaght	Linked hospital
Cavan General Hospital	Linked hospital
Connolly Hospital, Blanchardstown, Dublin	Linked hospital
Kerry General Hospital	Linked hospital
Letterkenny General Hospital	Linked hospital
Mayo General Hospital, Castlebar	Linked hospital
Mercy Hospital, Cork	Linked hospital
Midland Regional Hospital, Tullamore	Linked hospital
Naas General Hospital, Kildare	Linked hospital
Our Lady of Lourdes Hospital, Drogheda	Linked hospital
Portiuncla Hospital, Ballinasloe	Linked hospital
Sligo General Hospital	Linked hospital
South Infirmary Hospital, Cork	Linked hospital
St. Joseph's Hospital, Clonmel	Linked hospital
St. Luke's Hospital, Kilkenny	Linked hospital
St. Luke's Hospital, Rathgar, Dublin	Linked hospital
Wexford General Hospital	Linked hospital
Our Lady's Children's Hospital, Crumlin, Dublin	Paediatric centre/Classified as a linked hospital for the purpose of this research

Appendix 14. Proposed interview topics



Proposed Interview Topics

Interviewer contact details: Patricia Heckmann (heckmanp@tcd.ie): 087 7690161

Data collection/use in cancer control

1. Can you tell me what you perceive to be the main benefits of collecting data for cancer control?
2. Which agencies are you aware of that collect data on cancer for the purpose of cancer control or performance management?
3. Are you aware if the reporting of these data is mandatory or optional?
4. Which standards are you aware of that are utilised in the current data collection for cancer control?

Questions – Challenges to the Secondary use of health data

5. Can you tell me what you perceive to be the main "*challenges*" to the secondary use[i] of data in the area of cancer control? A number of subheadings have been included below as themes.
    a) Inefficiency
    b) Variations in measurement systems
    c) Organisational and cultural issues
    d) Technological barriers for electronic health records
    e) Economic pressures
    f) Competing priorities

Questions – Systemic Anti-Cancer Therapy (SACT)[ii]

6. Why do you think that the SACT reimbursement scheme has been introduced?
7. Who do you perceive the stakeholders involved in this scheme to be?
8. Would you suggest other data that could be captured that might be beneficial?
9. Do you think that these data have a wider role within a national system?
10. What approach do you think could be taken to provide the missing data?
11. What do you perceive as the main challenges to the provision of these data?
12. What would act as facilitators to the provision of these data?

Questions – Conclusion

13. What steps could be taken to improve the secondary use of data?

---

[i] In 2007 The American Medical Informatics Association defined the secondary use of health data as "non-direct care use of personal health information including but not limited to analysis, research, quality/safety measurement, public health, payment, provider certification or accreditation, and marketing and other business including strictly commercial activities" (Safran et al., 2007). This encompasses reuse both within and between individual organisations and for a variety of purposes.

[ii] Traditionally, drugs administered in secondary care, including SACT, were funded as an integral part of each hospital's overall budget. In 2011, the NCCP proposed that the funding stream for secondary care SACT be disconnected from the overall secondary care budget and instead funded through direct reimbursement to the hospitals based on patients' consumption of these medications (HSE, 2012a).

On initial analysis it would appear that 90% of the required data resides within one hospital system, and the remaining data are available only on the paper chart or else in separate systems.

Appendix 15. Semi-structured interviews – notes record form

Hints to be given when not originally volunteered

Preamble:

As per our earlier correspondence you will be aware that I am a student currently undertaking a Trinity College Dublin Masters in Health Informatics. I am undertaking a research study as a component of my research study.

The purpose of the study is to demonstrate the challenges to the secondary use of health data in cancer control in Ireland. The study is utilising a recently implemented reimbursement web portal as a case study. The portal requires manual submission of data to facilitate reimbursement for systemic anti-cancer therapy (SACT) administered in public hospitals.

Interview with:

Interview venue:

Date:

Time:

Consent received: Written Verbal

Sections - data collection/use in cancer control

Sections - data collection/use in cancer control		**Hints (tick when answered)**	**Additional comments**
1.	Can you tell me what you perceive to be the benefits of collecting data for cancer control?	• Access to service • Improve outcomes • Statistic • Incidence • Guideline adherence • Protocols adherence	
2.	Which agencies are you aware of that collect data on cancer for the purpose of cancer control or performance management?	• NCCP • HIPE • NCRI • HIQA	
3.	Are you aware if the reporting of these data is mandatory or optional?	• NCCP • HIPE • NCRI • HIQA	
4.	Which standards are you aware of that are utilised in the current data collection for cancer control?	• HIQA SBD - datasets • HIPE codes • ICD • NSAI • ISO • NCRI dataset • NCCP datasets	

Sections – challenges secondary use of health data

In 2007 The American Medical Informatics Association defined the secondary use of health data as “non-direct care use of personal health information including but not limited to analysis, research, quality/safety measurement, public health, payment, provider certification or accreditation, and marketing and other business including strictly commercial activities” (Safran et al., 2007). This encompasses reuse both within and between individual organisations and for a variety of purposes.

Sections – Challenges Secondary use		**Hints (tick when answered)**	**Additional comments**
5.	Can you tell me what you perceive to be the main challenges to the secondary use of data in the area of cancer control?		
a.	Inefficiency	• Variations in data collection. - Collection and reporting utilise different taxonomies and data definitions leading to requirements for data validation and continuous updating • Documentation and data quality. (Availability of data) - Incomplete clinical documentation, - Disparate electronic systems, - Manual data abstraction, - Inconsistent policies and practice • Provider staff resources: - Increased staffing resources in conjunction with reporting requirements	
b.	Variations in measurement systems	• Mandatory vs. voluntary reporting • Differing reporting formats for differing institutions sometimes for the same disease and patient cohort	

Sections – Challenges Secondary use		Hints (tick when answered)	Additional comments
c.	Organisational and cultural issues	• Health care organisations must react to changing requirements. This requires the organisations to have stakeholder acceptance, internal change organisation and culture that will allow the continuing provision of reliable data.	
d.	Technological barriers for electronic health records (EHRs)	• Uncoordinated implementation of health IT systems locally and nationally • Interoperability issues • Cost • Lack of understanding of the improvement role that EHRs can play in improving data reporting nationally • Lack of minimum common data sets for population health and quality measurement • Concerns with regard to implementing EHRs with regard to - Security and privacy issues - Data ownership issues	
e.	Economic pressures	• Costs of collecting data • Cost of dissemination and interpretation of performance data within organisations	

Sections – Challenges Secondary use		**Hints (tick when answered)**	**Additional comments**
f.	Competing priorities	• Variations in measure sets, data metrics and taxonomies - Lack of alignment between the institutions mandating the reporting - Absence of a national health care quality data set and report card • Privacy of individuals versus reporting requirements (data stewardship) • Keeping up to date with the changing reporting requirements	

Sections – SACT reimbursement scheme		**Hints (tick when answered)**		**Additional comments**
Traditionally, drugs administered in secondary care, including SACT, were funded as an integral part of each hospital's overall budget. In 2011, the NCCP proposed that the funding stream for secondary care SACT be disconnected from the overall secondary care budget and instead funded through direct reimbursement to the hospitals based on patients' consumption of these medications (HSE, 2012b).				
6.	Why do you think that the SACT reimbursement scheme has been introduced?	• Consolidation/centralisation of budget • Clinical audit - EBM • Outcome analysis • Effectiveness research		
7.	Who do you perceive the stakeholders involved in this scheme to be?	• NCCP • NCRI • HIQA • Hospitals • Oncologists • Nurses • Pharmacists	• Financial/admin Patients • PCRS • DOH • HSE • Clinical researchers	

Sections – SACT reimbursement scheme		Hints (tick when answered)	Additional comments
8.	Do you think that these data have a wider role within a national system?	• NCCP • NCRI • HIQA • Personal Health Record • National Clinical Information System • National Cancer Information System	
9.	Would you suggest other data that could be captured that might be beneficial?	• Dose in mg/g • Height • Weight • ECOG • Tumour staging	
On initial analysis, it would appear that 90% of the required data resides within one hospital system, and the remaining data are available only on the paper chart or else in separate systems.			
10.	What approach do you think could be taken to provide the missing data? **90% available**	• Upgrade current systems • Change data set • National system	
11.	What do you perceive as the main challenges to the provision of these data?	• Disparate electronic systems • Paper charts • Incomplete data • No ownership • Lack of commitment • Data protection • Manual entry • Resource –staff • Resource – hardware/software	

Sections – SACT reimbursement scheme		**Hints (tick when answered)**	**Additional comments**
12.	What would act as facilitators to the provision of these data?	• Identifiers o Personal o Institutional o Providers o Drugs • Support o IT o Staff o Financial • Interfaces • Standard protocols	
13.	Any other comments		

	Sections – Conclusion	**Hints (tick when answered)**	**Additional comments**
14.	What steps could be taken to improve the secondary use of data?	• Minimum data sets • Implementation of standards • Identifiers o Personal o Institutional o Providers o Drugs • Framework • Reuse of existing data • Resources	
15.	Any other comments		

Appendix 16. Stakeholder details

Primary Stakeholder	**Stakeholder group**	**Domain area**	**Interest**
Y	NCCP	Cancer Control	NCCP
Y	PCRS	PCRS	PCRS
Y	Hospital user	Hospital Pharmacist – large centre Location of data	Hospital
Y	Hospital	Hospital Pharmacist -small centre Location of data	Hospital
Y	Hospital	Data input	Hospital CUH
Y	Hospital	Financial person	Hospital CUH
Y	NCCP	Financial requirements	Financial/Medical Oncology Programme NCCP
Y		Data requirements	Programme manager NCCP
Y	PCRS		PCRS ICT
N	HIQA	National standards for healthcare data and KPIs	HIQA
N	Cancer Registry	Cancer Registry and Cancer Control	NCRI
N	NCCP	Cancer Control, Cancer Intelligence, Public Health	NCCP
N	Hospital		Consultant Oncologist BH
N	HSE ICT	HSE ICT	HSE ICT
N	HSE ICT	HSE ICT	HSE ICT
N	Vendors	System supply Location of data	Vendors
N	Hospital	Data manager hospital based	Data manager MMUH
N	Hospital	Interfacing disparate systems	Interface specialist MMUH
Stakeholders – not interviewed			
N	Patients		Equity and access to treatment
N	Government	Legislation	DOH
N		Data protection	DPC
N	Hospital	Data integration EHR	EHR project manager MMUH
N	Drug companies	Drug supply	Senior key account managers