ESTABLISHMENT OF METHODS FOR EXTRACTING AND ANALYSING PATIENT DATA FROM ELECTRONIC PRACTICE MANAGEMENT SOFTWARE SYSTEMS USED IN FIRST OPINION VETERINARY PRACTICE IN THE UK

Thesis submitted to the University of Nottingham for the degree of

Doctor of Philosophy

By

Julie Susan Jones-Diette, BSc (HONS)

July 2014

Abstract

Examining patient records is a useful way to identify common conditions and treatment outcomes in veterinary practice and data gathered can be fed back to the profession to assist with clinical decision making. This research aimed to develop a method to extract clinical data from veterinary electronic patient records (EPRs) and to assess the value of the data extracted for use in practice-based research. The transfer of new research from continuing professional development (CPD) into practice was also considered.

An extensible mark-up language (XML) schema was designed to extract information from a veterinary EPR. The analysis of free text was performed using a content analysis program and a clinical terms dictionary was created to mine the extracted data. Data collected by direct observation was compared to the extracted data. A review of research published in the proceedings of a popular veterinary CPD event, British Small Animal Veterinary Association (BSAVA) Congress, was appraised for evidence quality.

All animal records were extracted and validation confirmed 100% accuracy. The content analysis produced results with a high specificity (100%) and the mined data analysis was successful in assessing the prevalence of a specific disease. On comparison, the data extracted from the EPR contained only 65% of all data recorded by direct observation. The review of BSAVA Congress abstracts found the majority of the clinical research abstracts (CRAs) presented to be case reports and case series, with differences in focus between CRAs and veterinary lecture stream abstracts.

This study has demonstrated that data extraction using an XML schema is a viable method for the capture of patient data from veterinary EPRs. The next step will be to understand the differences found between data collected by observation and extraction, and to investigate how research presented as CPD is received, appraised and applied by the veterinary profession.

“It’s never too late to be what you might have been”

George Eliot

English novelist (1819 - 1880)

Acknowledgements

Firstly I'd like to say thank you to my fellow team members in the Centre for Evidence-based Veterinary Medicine and my fellow post grads in the School of Veterinary Medicine and Science, both past and present. You have encouraged me and stimulated discussion and thought over the past 3 years offering support and particularly humour when needed. I will cherish the relationships I have formed and the memories of my time as a post grad. Thank you to my supervisors Rachel Dean, Marnie Brennan and Malcolm Cobb who have spent so much time with me throughout my PhD listening and advising on all aspects of the work. They have been a constant source of support and encouragement. Thank you to Tim Parkin and Julia Yates for guidance on content analysis and Tracey Hassall-Jones for help with database design. Finally thank you to the PMS system provider and their clients for their collaboration and assistance, also to The Centre for Evidence-based Veterinary Medicine, The University of Nottingham and Novartis Animal Health for sponsoring this research.

To my lifelong friends, who never really know what I do but offer understanding anyway and have patiently waited for me to finish writing before I can resume my social activities, I thank you all. I am excited to get my other life back. To my family, both here and overseas, I am so lucky to have such wonderful people in my life. To my brother Kevin who protected me when I wasn't big enough to stand up myself against the horrible boys next door and was my hero for that, my sister Lesley who taught me how to be strong and stand up for myself, who made me so determined and with whom I love to laugh and finally to my wonderful Mum and Dad who somehow kept us all in line, safe and very very happy. I couldn't have achieved so much without their love and caring.

To my pets, they have sat, slept and played by my side while I worked, they have kept me company on the long nights of writing and I love them all so much.

Finally to my husband Mike I dedicate this work. The most encouraging and generous husband I could wish for, patient and caring, you gave me the opportunity, time and support to follow my dreams and I will be eternally grateful.

In memory of my very special Dad

I miss you every day.

Although you are gone, I know I never walk alone.

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter One

1 Review of the Literature

1.1 Introduction

Evidence-based Medicine (EBM) is the use of the best relevant, most up-to-date information in clinical practice to enhance clinical decision making and improve patient care (Sackett DL, 2000, Cockcroft and Holmes, 2003, Faunt et al., 2007, Forrester and Roudebush, 2007).

The importance of scientific and rigorously tested research to clinical practice was suggested as far back as the 1970s, when Archie Cochrane published his thoughts on how human healthcare could be improved (Cochrane, 1972). However the move towards evidence-based clinical decision making has only become popular in human healthcare since the 1990s (Evidence-based working group, 1992, Sackett, 1995, Sackett et al., 1996) and the veterinary profession has also been quite slow to respond, waiting a further decade.

The earliest publication to discuss evidence-based veterinary medicine (EVM) specifically, according to Margaret Slater (2010), appears to be a chapter by Bonnett and Reid-Smith published in the Vet Clinics of North America in 1996 (Bonnett and Reid-Smith, 1996). In 1998 an article by Malycincz appeared in the Veterinary Record discussing Evidence-based veterinary medicine (Malynicz, 1998) and in 2003 the first textbook was published by Cockcroft and Holmes entitled Handbook of Evidence-based Veterinary Medicine (Cockcroft and Holmes, 2003).

Veterinarians wishing to practice EVM currently may be limited by many factors; one obstacle being the small number of high quality peer reviewed clinical trials and published research papers (Muir, 2003, Vandeweerd et al., 2012a, Vandeweerd et al., 2012c). Evidence-based veterinary medicine can only be employed by practicing veterinarians when relevant evidence exists, and for some areas of veterinary science there is little evidence available (Holmes, 2007, Vandeweerd et al., 2012c).

A lack of available research and limited opportunity in veterinary practice to perform research due to time or constraints of cost or due to ethical considerations, has contributed to the poor level of available applicable or current evidence (Kapatkin, 2007). In her paper on outcome-based clinical practice, Kapatkin (2007) claims the answer to be “multicentre research using standardised study design and outcome measures where information can be synthesised together to create a greater body of evidence cheaply and quickly to support evidence-based medicine”.

This is where Evidence-based Veterinary Medicine (EVM) and Veterinary Epidemiology are most aligned, as stated by Margaret Slater in her review of both disciplines “...epidemiology studies populations and (evidence-based veterinary medicine) is about taking population data and applying it to individual patients.....Epidemiology is the cornerstone of EBM. Epidemiology can provide the high quality evidence needed for veterinary medicine” (Slater, 2010).

1.2 Veterinary Epidemiology and Evidence-based Veterinary Medicine

1.2.1 What is evidence?

Evidence-based medicine (EBM) and evidence-based veterinary medicine (EVM) have been defined in many different ways but generally are described as the use of the best available evidence generated from well-designed systematic research combined with clinical expertise and a consideration of the circumstances of the patient or client in making optimal clinical decisions (Sackett DL, 2000, Cockcroft and Holmes, 2003, Slater, 2010).

So what is the best evidence available? It is information collected from well-designed high quality research (Slater, 2010, Greenhalgh, 2010), it is (occasionally) peer reviewed and often described using a hierarchy of strength such as that presented in Figure 1-1 below.

Systematic Review
Meta-analysis

Experimental
Randomised Controlled Trial (RCT)
Controlled or Randomised Clinical Trial

Observational
Cohort
Case Control
Cross Sectional study or Survey

Descriptive
Case Series
Single case report

In-vitro
Clinical trial
Diagnostic methods
Unclassifiable/ Discovery research

Figure 1-1 List of study type presented in the classic hierarchy of evidence for intervention research according to (Kastelic, 2006, Lean et al., 2009).

The information provided by the differing study types for answering a clinical question depends on how well the study is designed and if the study design is appropriate for the aims. Experimental studies such as randomised controlled trials (RCTs), if well designed, offer the most powerful evidence to measure treatment effects, as randomisation reduces the risk of bias allowing greater confidence in the outcome of an effect being a result of the treatment given.

Observational studies, such as cohort trials, also have value particularly for identifying risk factor associations. They are cheaper to run than RCTs especially retrospective cohort studies where recruitment of patients is not required (Cockcroft and Holmes, 2003, Vandeweerd et al., 2012b).

The different types of study design and their relative strength has been explained fully by Kastelic (2006) and also by Greenhalgh (2010) who have both written excellent reviews describing in detail the different study types and level of evidence provided by each.

1.2.2 The role of Epidemiology in Evidence-based Veterinary Medicine

So what is the relationship between veterinary clinical epidemiology and evidence-based veterinary medicine (EVM)? Veterinary epidemiological studies can provide the necessary evidence needed to support an EVM way of working (Slater, 2010). Evidence-based veterinary medicine is at the clinically applied end of the spectrum of epidemiology and well-designed clinical epidemiology studies can provide much needed information to support veterinarians in practice and clinical decision making in the form of research results.

Clinical epidemiologists have the necessary skills and expertise to synthesise information collected from practice-based research into useable and easily understood research for the busy practitioner. Time saving tools from clinical epidemiology are now emerging to support evidence based methods in veterinary medicine such as critically appraised topics (CATs); a focussed and succinct review intended to answer a specific clinical question (Cockcroft and

Holmes, 2003, Faunt et al., 2007, Shearer, 2013) and also refer to the following websites for further information; www.nottingham.ac.uk/cevm, www.bestbets.org, www.bestbetsforvets.org, www.cebm.net. Critically appraised topics are perfect for use in clinical practice and can be a great teaching tool (Slater, 2010). They are prepared in a readily available and explained format, perfect for use by a busy practitioner who may not have the time to perform the search and retrieve exercise required to assimilate the evidence they need (Holmes, 2005, Slater, 2010). Although sometimes criticised for their transient value as the information they provide can quickly go out of date, in a profession where little evidence-based decision support exists they provide the best most current evidence available (Slater, 2010).

1.3 Veterinary Epidemiology and Practice-based Research

In medicine it is accepted that there is a need for clinicians to understand the principles of basic clinical epidemiology to support an EBM way of working (McAllister and Wild, 2009). Although this is something not always easily applied without specific training, there is value in the combination of veterinary epidemiology and evidence-based veterinary medicine in first opinion veterinary practice (Cockcroft and Holmes, 2003). The next step toward evidence-based veterinary practice would be to identify key areas in which research is needed (Figure 1-2). In order to identify where there are current gaps in knowledge for the clinical decisions veterinary practitioners are faced with, it would be useful to understand which clinical conditions are least understood or which are presented most often.

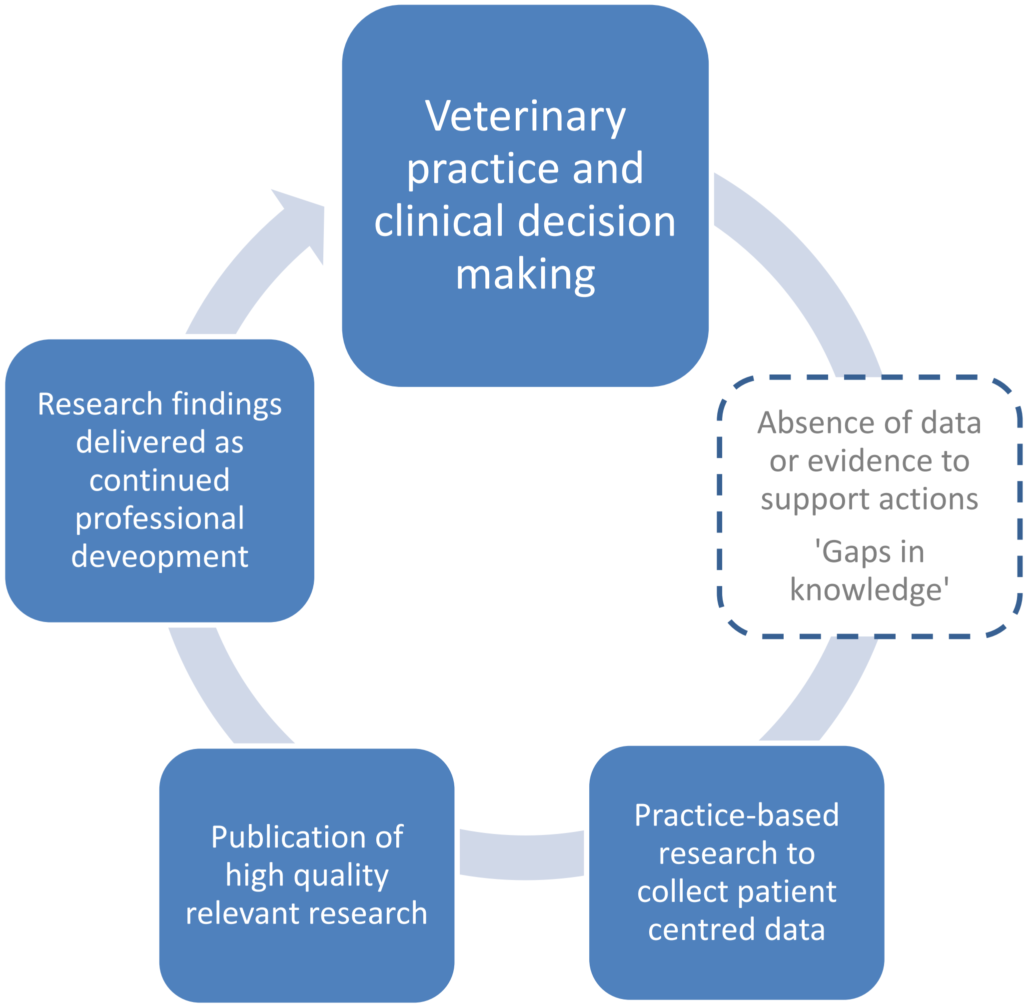


Figure 1-2 The cycle of practice-based research and CPD in evidence-based veterinary medicine and clinical decision making. Highlighting the missing link – Gaps in knowledge.

Information gathered from patient records have been used successfully for clinical epidemiology in the medical field, in particular the collation of data collected from first opinion practice, and research examining clinical outcome in human medicine has made important contributions towards the effective transfer of research findings into clinical practice (Whitelaw et al., 1996, Richard D Neal et al., 1996, Walley and Mantgani, 1997, Rodriguez and Gutthann, 1998, Hippisley-Cox et al., 1998, Thiru et al., 1999, Langman et al., 2000, Black et al., 2002, Hammersley et al., 2002, Susan S. Jick et al., 2003, Hippisley-Cox and Stables, 2011). Practice-based research is a valuable part of this process and in particular, the aggregation of patient outcome data has been of great use to human health epidemiology (Hammersley et al., 2002, Holt et al., 2008, Hippisley-Cox and Stables, 2011).

Examining patient records as a population is a useful way to identify common conditions and treatment outcomes seen in veterinary clinical practice (Faunt et al., 2007) and data gathered can be fed back to the profession to assist with clinical decision making. It is now recognised that practice management software (PMS) systems hold a wealth of clinical information (Faunt et al., 2007). Utilising medical informatics methods (Szolovits, 2003, Uzuner et al., 2006, Meystre et al., 2008) to access veterinary electronic data is becoming increasingly popular for veterinary clinical epidemiology and current methods focus primarily on extracting and analysing the information held within electronic patient records and clinical notes (Lund, 1997, Estberg et al., 1998, Faunt et al., 2007, Lam et al., 2007a, Moore et al., 2007, Anholt, 2013).

The application of medical informatics methods to extract and analyse electronic patient records (EPRs) from first opinion veterinary practice will hopefully open up a wealth of opportunities for practice-based research and support the move towards more evidence-based veterinary practice.

1.4 Clinical Epidemiology and Informatics

The examination of patient records and the aggregation of patient data for clinical epidemiological research have been used effectively in human healthcare in the United Kingdom for the analysis of disease prevalence and incidence in human populations (Walley and Mantgani, 1997, Hippisley-Cox et al., 1998, Thiru et al., 1999, Hammersley et al., 2002, Holt et al., 2008). Methodologies such as this have increased in popularity over the last 10 years with the introduction of computerisation for practice records.

The move from paper to computerised record systems for clinical record keeping within veterinary first opinion practice has created opportunities for veterinary clinical epidemiology (Summers et al., 2010, Radford et al., 2010, Diesel et al., 2010, Mateus et al., 2011), and much can be gained by looking to the advances made by medical informatics in the last decade. Human healthcare research methodologies such as those used by Q Research at the University of Nottingham, where GP patient data is collected and combined for population based research, can provide an excellent model for veterinary informatics and practice-based research (Hippisley-Cox and Stables, 2011).

Practice-based research and the aggregation of patient data from many vet practices in addition to a standardisation of research methods would facilitate the multicentre philosophy recommended by Kapatkin (2007). Little information is currently available on the prevalence and incidence of disease in the companion animal population, with some exceptions (Radford et al., 2010, O'Neill et al., 2012, Anholt, 2013), and more data is needed from clinically orientated basic research in first opinion veterinary practice. Data collected from PMS systems in first opinion practice would allow an understanding of the challenges faced by vets each day and also provide veterinarians with the information to help guide clients to choose the best treatment possible.

1.5 Medical Research

1.5.1 The Electronic Patient Record

Since the 1970s there has been an increasing awareness of the need for evidence of clinical effectiveness and outcomes research in human healthcare (Cochrane, 1972, Tunis et al., 2003). Outcomes research considers which treatment or interventions work best under certain circumstances and provide evidence to support this. Outcomes research is favoured by the UK National Health Service (NHS) who for some time has been made to account for the care they provide due to failings in care and increases in the cost of providing healthcare (Tunis et al., 2003).

A decade ago the variability in clinical practice and outcome between healthcare providers led to a call by leaders of clinical governance for a more evidence-based approach to human medicine and a need for more high quality evidence (Hassey et al., 2001, Tunis et al., 2003). This has resulted in the development of methods to provide reliable and comprehensive information to both the healthcare profession and key decision makers. Outcome research has been an important part of this transition and clinical research now has greater value for translating knowledge into clinical practice (Tunis et al., 2003).

In an attempt to address the gaps in knowledge, the medical records of both primary and secondary healthcare providers were considered an excellent source of high quality health information (Hassey et al., 2001). However it was a task hindered by the format of the clinical notes. Handwritten, varying in content and structure and with no governance over organisation; the patient record was not an easy body of information to access (Ferranti et al., 2006). By the 1980s there was an increased need for well structured and readily accessible patient information which coincided with advances in computer science and for the next 25 years the electronic patient record (EPR) was developed and refined (van Bemmel and Musen, 1997).

The structure and organisation of hospitals and secondary healthcare made the introduction of electronic systems an easier task. However by the 1990s computerised healthcare records became more readily accepted by primary healthcare and by 1997 at least 50% of GP practice records in the UK were electronic (van Bemmel and Musen, 1997).

As medical informatics has evolved, the computer in clinical practice has become more important as a tool for organisation and research (van Bemmel and Musen, 1997, Shortcliffe and Blois, 2003, Chen et al., 2005). A study by Rodnick (1988) surprisingly found only 5% of GPs in the United States were using their practice computer for more than just billing. Computers in general practice in the UK are now used for a number of different functions from recording and reporting test results, access to scanned or transcribed reports, and providing information to assist with decision support, administration and finance, and finally research for measuring treatment or procedural outcomes, clinical audit, benchmarking and supporting clinical trials (Shortcliffe and Blois, 2003).

1.5.2 Information Extraction

In June 1987 the UK General Practice Research Database (GPRD) was set up (Walley and Mantgani, 1997). This was a commercially managed research databank created with the intention of linking patient records, including signalment data, with the prescription of drugs and any associated co-morbidity, co-prescribing, adverse effects or reaction to drugs.

General practitioners (GPs) were encouraged to participate in the project by being provided with a free computer for their clinical records. The computer contained an integral practice management software system, created by VAMP Computer systems, training and support in turn for participation in the project (Lis and Mann, 1995).

Due to the unique nature of the UK National Health Service (NHS), all UK patients are offered healthcare and are required to be registered with a general practitioner who is responsible for the majority of their healthcare needs. As such the health records for the UK population are very complete and comprehensive. The GPRD study ran until 1994 and estimates made at that time suggest 98% of the UK population was registered with a GP practice.

The GPRD proved to be very successful and by 1993 the databank held information from 4.4 million patients registered with 680 practices and 2333 individual GPs (Lis and Mann, 1995). The database included over 10 million patient-years of records and set up costs between 1987 and 1993 were reported to be £14 million. Unfortunately the GPRD suffered from its own success and proved to be a very costly project to manage; as a result it was eventually interrupted due to lack of funding and government support.

Since 1994 the Secretary of State for Health has owned the GPRD database which is now managed by the Medicines Control Agency – MCA which became part of the newly created Medicines and Healthcare products Regulatory Agency (MHRA) in April 2003. Recently relabelled as the Clinical Practice Research Datalink (CPRD), the MHRA have picked up the responsibility for the data collection and CPRD are now the new observational data and research service for the English National Health Service (NHS) and is integrated with the MHRA. This joint initiative is funded by the NHS, National Institute for Health Research (NIHR) and the Medicines and Healthcare Products Regulatory Agency (MHRA)[1]. Since the creation of the GPRD many other groups have followed their lead.

Developed to offer a more systematic approach to the detection of disease incidence and risk in general practice, the NHS 'general practice networks in primary care' initiative has been created (Mant, 1997). This group of databanks combine data from individual first opinion healthcare providers with other practices within a centrally managed hub for the purpose of clinical data analysis and measures of treatment outcome. Participating research groups include; The Primary care Information Services project (PRIMIS+) (www.primis.nottingham.ac.uk/), Q Research (www.QResearch.org), Clinical Practice Research Datalink (CPRD) (www.cprd.com) and The Health Improvement Network (THIN) (www.thin-uk.com).

To participate in the research a GP practice must have a PMS system which is a member of the group initiative. Once a GP practice has joined the network,

[1] www.cprd.com

the hub will run language queries on the practice management software system background database, collating clinical data at an individual practice level. These systems are also able to combine data from further afield at the local authority level, via primary care trusts, creating a network of information retrieval and sharing. This data can then be accessed by governing bodies, such as those at a national level, to create published guidelines for patient treatment; for example those published by the National Institute for Health and Clinical Excellence (NICE)[2].

The MIQUEST Project of the Health and Social Care Information Centre (HSCIC) (www.hscic.gov.uk/systems) has developed bespoke data extraction software, called MIQUEST Enquirer, to assist those in the NHS responsible for the collection and management of patient data in the NHS (De Lusignan S et al., 2006), and QResearch (Hippisley-Cox et al., 1998, Hammersley et al., 2002, Tim A Holt et al., 2008) and the Royal College of General Practitioners Disease Surveillance Group (RCGP) (Scherer R et al., 2007, Crombie, 2010) have all provided and published epidemiological data to support NHS research.

[2] www.nice.org.uk.

Recently Q Research and THIN have also developed patient-level aggregated databases for the purpose of running research queries for human clinical epidemiology research. These systems use extracted coded data from millions of patient encounters (Smith et al., 2008, Hippisley-Cox and Stables, 2011) and are now also considering the analysis of free text data. Database extraction and analysis of this type provides data for both longitudinal and cross-sectional studies, they offer a unique source of population-based information and have proved to be valuable for epidemiology research within the NHS (Lawrenson et al., 1999).

Finally in Great Britain the PRIMIS+ project was designed to assist the NHS to improve their patient care. The system is by GP subscription and works by querying clinical data in the EPR using the MIQUEST query language to provide primary care quality measures, training and benchmarking data (Hassey et al., 2001).

Most of the progress made in medical informatics has been due to the extraction of information from the electronic patient record (EPR). However the move from paper to electronic records has been relatively slow and many hospitals still used paper for their emergency care, in-patient and discharge summaries (Hersh, 2003a).

1.5.3 Data Quality of the Electronic Patient Record

General practice EPRs in the UK are unique in that they are expected to follow a single patient for their entire lifetime, including when the patient moves - the record must follow and therefore needs to be compatible with many different GP's practice management systems. Data is usually entered quickly and at a high volume, the database must support many different conditions or symptoms from all ages, sexes and situations and be flexible to allow for the addition or transfer of information between multiple care providers (Bernstein et al., 1993).

As already mentioned, EPRs also now have value as a tool for research and as such new and innovative ways to access this wealth of diverse clinical information is required. However there are some limitations within the software and hardware for recording all the desired information and then extracting it. In addition there are certain challenges with the data itself. Data quality can vary and common causes for information error in the EPR reported by Bernstein et al (1993) are,

i. Human error, physicians entering information incorrectly or not at all.
ii. Laboratory or practice personnel errors during the input or transcription of results or records.
iii. Misinformation from patients.

iv. Concerns over loss of confidentiality restricting full accounts of events.

v. In addition GPs find the entry of data into some computer database systems restrictive and cumbersome to perform reducing their compliance to fully complete the record.

1.5.4 Standardised Medical Terminology

1.5.4.1 Clinical coding

The analysis of medical information extracted from the EPR also requires knowledge of linguistics in particular medical language and the semantics and synonyms of medical terms. Take this example adapted from Van Bemmel and Musen (1997) :

- Stomach is an organ
- Cough is a symptom
- AIDS is a disease
- Dyspnoea is a clinical sign which may also be referred to as shortness of breath.
- Pain requires description such as location, intensity, severity, progression, radiation.

All of this information needs to be accessed and organised for analysis in a unified way to understand its meaning and context within the record (Dolin, 1997). In addition the information can be presented in the negative e.g. No cough present, absence of pain.

To make full use of information recorded within an EPR a researcher must have a thorough understanding of the medical terminology and language used. Clinical coding and the standardisation of terminology can assist with this by classifying all possible variations of terms and medical dialogue to be allocated within a singular hierarchical system.

The introduction of nomenclature and clinical coding for human health language such as the NHS Read Codes (De Lusignan S et al., 2004), The International Statistical Classification of Diseases and Related Health Problems (ICD) codes (http://systems.hscic.gov.uk/data/clinicalcoding), and The Systematized Nomenclature of Medicine (SNOMED) (Rothwell et al., 1993, De Lusignan S et al., 2004), has been a significant support to medical informatics; particularly to identify records of clinical signs, diagnosis or presenting complaint. However it has to be assumed the codes are used appropriately and accurately for there to be value in the short cut they provide (De Lusignan S et al., 2004), and some research suggests this may not always be the case and that there is often little agreement between coded information and the free text within the same record (Stein et al., 2000).

1.5.4.2 Free text

There is a wealth of information available within the EPR; however as already discussed it is not always readily accessible. As medical informatics has gained in popularity over the last decade the focus has been on extraction of information from the EPR using automated systems. It has however become quite apparent that a large proportion of useful information is in the form of clinical narrative recorded during a consultation usually termed ‘free text’ (Hersh et al., 1997, Uzuner et al., 2006).

This type of information is not easily accessed or analysed by automated processes as it is usually unstructured, ungrammatical, fragmented typed notes (Spyns, 1996, Uzuner et al., 2006). Where records are predominantly free text there is often little consistency in the terminology or language used, making the task an even longer process (Meystre et al., 2008). Nevertheless extracting and analysing the free text is considered valuable for both clinical care and research as it holds information much richer than that contained within the diagnostic codes alone (Jollis et al., 1993, Hersh et al., 1997, Dolin, 1997) so the development of methods to access this information is essential.

1.5.5 Medical Informatics

Advances in medical informatics have more recently focussed on ‘unlocking’ all information within the clinical records using artificial intelligence and computerised linguistic techniques with some success, and research is ongoing (Szolovits, 2003, Uzuner et al., 2006). This has opened up the opportunity to examine the free text within the primary care EPR and also within some secondary healthcare records, such as discharge summaries. These records contain patient information not necessarily recorded in the primary care EPR including past medical history, co-morbidities and discharge diagnosis.

A popular method for the extraction and analysis of information from clinical records is Natural Language Processing (NLP); a field of informatics, artificial intelligence and linguistics which applies an automated system of analysis on electronically recorded human (natural) language in an attempt to produce meaningful interpretation and an output of aggregated and organised information.

Extensible mark-up language, (XML) is a computer language that can assist with encoding of documents in a format that is readable to both humans and computers. It is popular for data exchange because it can be used between systems and is simple and flexible using common language and terminology (Katehakis et al., 2001, Quin, 2012).

The use of XML language has opened up opportunities for the use of NLP processes by allowing the cross communication of systems. NLP in its simplest form is the extraction of information from natural language electronically using a list of search terms. However, much more sophisticated methods are becoming popular, in particular i2b2 (Informatics for integrating Biology to the Bedside), a project funded by the National Library of Medicine, i2b2 works by extracting and mapping information of interest to a standard structure for analysis which will hopefully remove the need for extra processing and organising of extracted information (Uzuner et al., 2006).

There are however limitations with medical informatics using the EPR. Recent research (Harman et al., 2012) has suggested information may actually be limited by the introduction of computerisation and EPRs into private practice. It is possible the combination of coding, data entry requirements and limited consultation time restricts the amount of information which can be captured per patient and information may be missed.

In a recent study of the medical profession in the US, researchers at the University of Florida (Harman et al., 2012) found that information was lacking, particularly in relation to mental health, from primary care consultation records when more than two chronic clinical conditions had been discussed during the consultation. It has been suggested that the relationship between GP and patient may be disrupted due to the necessary 'GP – computer' interaction where attention is given more to recording of information than observational identification of disease or signs of illness and this may be compounded when general practitioners are required to record multiple signs into separate note fields within an EPR (Harman et al., 2012).

In summary, the main purpose of the computer system within a GP practice is for patient care and as such the data is not easily controlled or research specific. Extraction of the data can be a complicated process and running queries can be repetitive and time consuming. The main barriers to utilisation appear to be time, expertise and money (Lawrenson et al., 1999). The systems are costly to maintain and require a lot of man power to manage and a high skill set for operators to run the analysis and data collation. The PMS systems are rarely designed or used in a way that allows easy access to the data required (Bernstein et al., 1993, Lawrenson et al., 1999, Ferranti et al., 2006). Additionally this type of research is difficult to validate (Hassey et al., 2001) and, if extracted, information needs to be cleaned or processed in some way which can be a very labour intensive exercise (Whitelaw et al., 1996, Hassey et al., 2001). However if successful, methodologies to extract and analysis this data offer a wealth of information for practice-based research.

1.6 Veterinary Research

The UK pet population is ever changing and growing and is difficult to quantify due to limited research on population dynamics, although see Downes *et al* (2013). Of the owned pet population, it is assumed there will be vet visiting and non-vet visiting pets. In addition the UK pet population is made up of owned and un-owned pets, such as those in rescue centres and shelters (Stavisky et al., 2012).

The numbers of animals within each group is unknown although work has been published which attempts to identify these figures (Downes et al., 2009, Downes et al., 2011, Stavisky et al., 2012, Murray et al., 2010). The Pet Foods Manufacturing Association (PMFA) Annual report for 2012 based on market research states that the total number of pets owned in the UK to stand at 8 million dogs and 8 million cats. In the UK today PFMA research statistics suggest approximately 48% of households (13 million) own a pet, excluding fish. The exact method for quantifying the PFMA figures, unfortunately, is unpublished. However the report states the data was collected qualitatively via face-to-face interviews with 2,159 adults, aged 16+ using a nationally representative of households in UK conducted by a commercial research company (http://www.pfma.org.uk).

Pet care is a costly business for an owner and, in the current economic climate, limited funds and an increase in access to different types of electronic medical information mean that veterinarians are expected to explain their treatment decisions and associated costs. This makes the work of the veterinarian increasingly challenging (Neuberger, 2000, Coe et al., 2007) and a veterinarian must feel confident that they are able to offer the best care available and they have a responsibility to themselves and their clients, to not only be aware of the most recent research and results in veterinary medicine (Cockcroft and Holmes, 2003) but also to contribute to this knowledge base. This is easily achieved by maintaining comprehensive patient records and especially, where possible, detailed accounts of clinical signs, diagnosis, treatment and/or outcome. The added risk of owner-vet disputes, and with emerging initiatives such as a clinical audit within the veterinary practice, detailed and comprehensive records of client-animal-veterinarian interaction are essential (O'Connell and Bonvicini, 2007, Vandeweerd et al., 2012b).

1.6.1 The Paperless Practice

In 1991 the American Veterinary Medical Association (AVMA) authorised their committee on veterinary medical informatics to conduct a national survey on computer use in private clinical practice and level of interest in computer information service networks (AVMA Committee on Informatics, 1992). Four thousand three hundred and forty six vets were surveyed and 45% replied (1,945). At that time 58% of private clinical practitioners, exclusively treating small animals, had office computers. More recently in 2006 a survey was conducted by the Royal College of Veterinary Surgeons (RCVS) in the UK (Robinson and Hooker, 2006) which indicated as many as 94% of veterinary practitioners use a computer for their practice records.

The progression from paper to computerised practice records, once introduced, was a rapid transition and appears to have been accepted readily by the practicing veterinarian. In some instances practice records may have been converted from existing paper records, and some practitioners have developed their own system of electronic records while others have purchased a dedicated practice management system (Foley et al., 2005).

A patient's record should be a full account of every interaction they and their owner have had with the vet practice. It is therefore vital that this information is as detailed as possible (Faunt et al., 2007, Baker, 2008, Marshall, 2008c, Marshall, 2008a, Marshall, 2008b) and the computerisation of practice records claims to support the fast and secure storage of information.

The introduction of computers to veterinary practice has streamlined the consultation process somewhat by assisting the process of billing, ordering, stock control, sample analysis and patient records (Foley et al., 2005). However the conversion has generally been to support veterinary practice, as one would expect, and not necessarily practice-based research.

The provision, by PMS system providers, of data language and input fields which allow for the electronic transfer of information have helped developed the field of practice-based research for veterinary medicine (Lund et al., 1999, Moore et al., 2005, Lund et al., 2006, Faunt et al., 2007, Moore et al., 2007, Moore and Lund, 2009, Glickman et al., 2009, Bartlett et al., 2010, Glickman et al., 2011). Banfield® The Pet Hospital and their colleagues in the United States, were the first to exploit this opportunity in small animal practice under the direction of Dr Elizabeth Lund who first published her PhD thesis on diagnostic surveillance in companion animal practice in 1997 (Lund, 1997).

PetCHAMP, a program developed at the University of Minnesota, was used by Dr Lund for surveillance in companion animals seen in private veterinary practice and was one of the first surveillance initiatives (Lund, 1997). More recently the Small Animal Veterinary Surveillance Network (SAVSNET) was developed at the University of Liverpool. SAVSNET, in collaboration with VetSolutions PMS system and veterinary diagnostic laboratories, has allowed for disease specific surveillance in first opinion veterinary practice with the introduction of surveillance surveys into the consultation (Radford et al., 2010). In addition clinical coding of diagnosis or clinical signs have allowed for the introduction of searchable terms (Lund, 1997, Summers et al., 2010, Mateus et al., 2011)

Veterinary hospitals, often based in Universities, and larger veterinary franchise groups such as Banfield® which has clinics all over the US, have started to organise their own data into a more standard format to allow them to feedback the findings into their care of patients (Lund et al., 1999, Lund et al., 2005, Moore et al., 2005, Lund et al., 2006, Faunt et al., 2007, Lund, 2008a, Lund, 2008b, Moore and Lund, 2009, Glickman et al., 2009, Glickman et al., 2011). Banfield has also examined data collected over the past 20 years through their own PMS system (Faunt et al., 2007) and have published findings for their veterinary surgeons (Banfield Journal) and the wider community.

The Veterinary Medical Database (VMDB) was originally developed by the National Cancer Institute for cancer surveillance in 1964 (Bartlett et al., 2010) and is managed by the University of Illinois. This database collects data submitted by contributing veterinary schools and assimilates this information into a useable database which is then available as a research tool (Bartlett et al., 2010).

Finally the Veterinary Medical Teaching Hospital (VMTH), University California Davis attempted to develop their own computerised patient record 30 years ago. However by 2001 little progress had been made apart from a paper published in 1998 on the extraction of veterinary information collected from the hospital information system using free text analysis (Estberg et al., 1998).

The advances in veterinary research discussed here, although significant, have experienced the same barriers as discussed previously in human healthcare research; namely a lack of standardised medical recording and limited medical terminology standards. In addition there is currently little opportunity to use clinical coding in veterinary medicine and the available information systems are not specifically designed to support research, making access to data a challenge.

1.6.2 The Standardisation of Veterinary Information

1.6.2.1 Clinical coding

The information recorded within the animals patient record by the veterinarian during a typical consultation is often recorded as conversational text, shorthand, acronyms, and synonyms and rarely appears consistent between veterinarians. McCurdy (2001) found PMS systems have a far greater value as tools for research than their current role in practice but that this value will not be realised until practice-based records are collecting similar information from each case using a standard nomenclature. Therefore the standardisation of veterinary terminology is considered to be necessary for the progression of practice-based research and disease surveillance (Slater and Boothe, 1995, Zelingher et al., 1995, Estberg et al., 1998, Anholt, 2013)

Efforts have been made to create a nomenclature for use in veterinary practice. VeNom (Veterinary Nomenclature) was developed by the Royal Veterinary College (RVC) (Diesel et al., 2010, O'Neill et al., 2012) and has been further developed and maintained by The VeNom Coding Group (www.venomcoding.org). The codes are based on the SNOMED coding system and are embedded into a PMS system within their own referral hospital at the RVC along with a small group of collaborating private veterinary practices. VetCompass (http://www.rvc.ac.uk/vetcompass/) is a research branch of the VeNom initiative and so is also co-ordinated by the RVC. VetCompass hope to use the clinical coding system VeNom for data collection of disease surveillance and breed specific disease. The current research uses a code which is chosen at the end of consultation by the practicing veterinarian to describe their diagnosis and, in some cases where a diagnosis cannot be confirmed, the clinical signs are coded along with species and breed information.

Other coding systems include SNOMED and the standard nomenclature of disease and operations (Human - SNDO and Veterinary SNVDO) which are used by the veterinary hospitals that contribute to the Veterinary Medical Database (VMDB). PetTerms, a coding system for diagnostic terms developed by Elizabeth Lund to support her PhD research in 1997, was based on the SNOMED terms but referenced animal rather than human specific terms (Lund, 1997, Dear et al., 2011).

Clinical coding has limitations with respect to the amount of data captured during consultation; one code assigned for diagnosis may not be the whole picture and may lead to underreporting (Lund, 1997, Lund et al., 1999). In addition it is often found that coding is not accepted by all practices or may not be used at every consultation or the desired code may not be readily available which may lead to losses of information (Bartlett et al., 2010).

1.6.2.2 Free text

The use of text mining and content analysis of free text within the veterinary EPR to reduce the labour required to search the text, can also be used to identify the terms used by vets in practice to record data. Whilst limits to the success of this currently lie in the lack of standardised veterinary terminology within the profession, free-text mining has been used successfully alongside clinical coding in the analysis of data from general practitioners in human healthcare (Walley and Mantgani, 1997, Rodriguez and Gutthann, 1998). It has also been applied for veterinary medicine in the thoroughbred horseracing and breeding industry (Lam et al., 2007a) and within first opinion and referral veterinary practices (Lund et al., 1999, Diesel et al., 2010, Bartlett et al., 2010). The use of emerging informatics tools such automated content analysis and text mining for the analysis of unstructured data may overcome the issues with free text (Szolovits, 2003, Uzuner et al., 2006) and allow the analysis of data removing current barriers to practice-based research (Hersh, 2003b).

1.6.3 Electronic Communication

Accessibility to electronic information and the ability for programs to work across different information management or operating systems is a major hurdle to information transfer and extraction in healthcare research for both human (Hersh, 2003b) and veterinary medicine (Faunt et al., 2007). A new initiative has been created termed the VetXML Consortium. Developed by Vet-Envoy, a veterinary communications company, and the Society of Practicing Veterinary Surgeons (SPVS), a not-for-profit organisation offering guidance to the veterinary profession, VetXML was formed in 2009. The purpose of the consortium was to create better electronic communication between veterinary practice and their service providers. Diagnostic laboratories, insurance providers, and practice management software system developers have all come together to create a dynamic system of paperless information transfer. It is hoped the establishment of the consortium may assist in the communication between veterinarians and their vendors, particularly PMS providers, to aid in the extraction of veterinary clinical information from the PMS systems in veterinary practice for the purpose of research.

Vet-Envoy (http://www.vetenvoy.com) an information management group acts as a hub for members of the consortium to co-ordinate all exchange of information. This process is intended to reduce the time needed to complete and send paper forms and should improve not only practice-to-practice exchange of information but communication between surgery and out of hours emergency services, practice and laboratory for the analysis of samples, to insurance companies to allow the secure and efficient transfer of patient data for claims and many other opportunities for data transfer including practice-based research.

1.6.4 Challenges for practice-based research using data from the EPR

There are some problems to overcome before data in the electronic patient record within veterinary practice is readily available. Often in human healthcare the records cover the whole life of the patient. In veterinary practice the animal may move location and a large proportion of the record be lost or be paper-based prior to a move to an electronic system, the animal may die with no record of the event, the veterinarian may not even be advised. The consultation time is often limited to 10 minutes (Gray and Cripps, 2005) but may take longer so the records, often completed at the end of the consultation, may be completed in haste prior to the next patient arriving. Finally the system must be adaptable for all possible events not only practice-based research.

Another difficulty of accessing the electronic patient record for the collection of data is confidentiality. Many veterinarians feel their patients data is confidential and therefore may not feel comfortable with the electronically transfer of information. Lack of control over where the information goes and how it may be used has been a stumbling block for the progress of electronic information extraction. A telephone survey conducted in 1997 suggested veterinarians were very keen to create a database for their medical records, although they had concerns over data protection and security (McCurdy, 2001).

A second limitation with this type of methodology, as previously discussed, may be that much of the medical and clinical information is captured as free text and there is a lack of continuity in language or medical terms used to record the information in the EPR (Slater and Boothe, 1995, Zelingher et al., 1995, Estberg et al., 1998). Therefore an innovative approach to data accessibility is needed and advanced informatics methodologies are required.

1.7 Conclusion

The overall aim of this research was to support clinical decision making in first opinion veterinary practice by providing a means to gather clinical patient data for veterinary clinical epidemiology and outcomes research.

Limited access to high quality research is a major hurdle to evidence-based veterinary medicine (Slater, 2010). The value of data extraction to veterinary practice-based research is in the access to information on disease presentation and incidence. A system which can identify and extract animal patient data from PMS systems and combine it into a data warehouse for population research would be an excellent tool for veterinary clinical epidemiology, providing an understanding of conditions commonly observed in companion animal practice. Data extracted will also allow a comparison of the quality and quantity of extracted veterinary clinical information to that collected by direct observation during the consultation providing a means to measure the value of extracted patient encounter and outcome data for practice-based research.

In conclusion, it is hoped this research will ultimately provide patient data for veterinary clinical epidemiology research and the information collected in practice will address current gaps in knowledge. In addition this will provide a valuable resource for the CEVM for evidence-based veterinary research and to provide a basis for the communication of research findings back into practice and further support clinical decision making in first opinion veterinary practice.

1.8 Research Aims

The experimental aim of the PhD research was to establish and validate a veterinary informatics method of extracting clinical information from Practice management software systems (PMSs) that could be used for EVM and thus inform future clinically orientated epidemiology studies. A secondary aim was to compare the content of an extracted electronic patient record (EPR) to information recorded during observed veterinary consultation to measure the precision and value of the data collection methodology.

As previously stated, limited access to high quality research is a major hurdle to evidence-based veterinary medicine. Therefore, in addition to the experimental research, a review was performed of the subject areas covered during the largest UK veterinary practice-based congress, British Small Animal

Veterinary (BSAVA) Congress. The aim of the review was to examine the quality of clinical research available at BSAVA and the information offered as CPD in the form of veterinary lectures.

A sample of research and clinical CPD published by the BSAVA in their proceedings over a ten year period (2001 – 2010) were considered. The specific aims of the work were;

a. To investigate the frequency of certain topic areas chosen for VS and CRA presentations over time.
b. To investigate the most popular type of study design used for clinical research presented as CRA over time from a sample of CRA.

1.9 Organisation of the thesis

The following chapters will introduce the method, results and discussion from a three year programme of research. The Thesis is structured as follows;

Chapter 1 is an introduction to the work and a review of the current literature. The chapters that follow are structured as individual manuscripts for each research study and are therefore presented with the following section headings; Introduction, Materials and Methods, Results, Discussion, Conclusion.

Chapter 2 details the design and validation of an XML schema as a tool for the extraction of clinical data from a veterinary PMS system. In addition, this chapter presents the validation of an Access database for the management, processing and safe storage of extracted patient data.

Chapter 3 presents a method for content analysis of the clinical free text extracted from the veterinary patient record, and compared the performance of the WordStat programs 'Keyword in Context' and 'Keyword Retrieval' as a tool for text mining of the clinical narrative in an EPR. Chapter 4 will demonstrate the use of a bespoke clinical terms dictionary as a method to identify clinical indicators of disease and consider the use of terminology and language in the veterinary clinical narrative.

Chapter 5 is the final research chapter and presents the findings of a comparison of data collected from first opinion veterinary practice via electronic patient record extraction against that collected by direct observation in real time during the consultation. The recording of clinical information in the patient record by veterinarians in practice was also considered and discussed.

Finally Chapter 6 is a review of the research presented at the British Small Animal Veterinary Association (BSAVA) 2000 – 2010. The content and quality of a sample of research presented at BSAVA was reviewed and the delivery of

CPD to the veterinary profession in the form of veterinary lectures was examined.

Chapter Two

2 Information Extraction

2.1 Introduction

Patient encounter and outcome data may provide answers to many questions veterinarians face daily and could be used to contribute to the clinical management of disease, for the assessment of treatment success and to consolidate details of diagnosis and disease progression (Lund et al., 1999, Lund et al., 2005, Lund et al., 2006, Diesel et al., 2010, Finn et al., 2010).

The increasing reliance on practice management software (PMS) systems within veterinary practice means much of the data collected during patient encounters is captured within an electronic record (Wise and Yang, 1992, Lam et al., 2007a, Robertson-Smith et al., 2010). A large proportion of data held within the EPR is thought to be recorded as clinical narrative usually referred to as 'free text' (Szolovits, 2003, Meystre et al., 2008). In human medicine, analysis of the free text is already making good progress in clinical epidemiology and research (Tim A Holt et al., 2008, Tate et al., 2011, Wang et al., 2012). If accessible, the electronic record would be of great value for veterinary epidemiology studies and practice-based research and to support evidence-based decision making (Hornof et al., 2001, Bakken, 2001).

Information from veterinary primary care consultations is commonly kept on many different commercial practice management software systems (Hornof et al., 2001, Robinson and Hooker, 2006, Robertson-Smith et al., 2010). Examples of different veterinary PMS systems are presented in Appendix 1. Individual queries for information can be written and run on each of these systems to enable the aggregation of patient data. However, due to differences in database structure between these systems, it is much more difficult to combine queries or requests for information across different systems although some efforts have been made towards greater accessibility and communication[3]. A versatile approach is vital if data is to be combined from many different veterinary practices utilising many different PMS systems.

To support clinical decision making it is necessary to understand which conditions present most commonly in small animal veterinary practice, identify diseases that are becoming more prevalent and understand more about case management. It is important to consider, in relation to the clinical workload, where there are relative gaps in current knowledge. By collecting

[3] www.vetxml.org

and collating patient data and generating hypothesis for research one can then work towards designing and conducting robust studies that will provide the high quality evidence needed most by practicing veterinarians.

The purpose of this research was to establish and validate a method for the extraction of small animal patient records from a veterinary practice management software system. A secondary aim was to ensure the information extracted was precise and the patient record could be extracted completely and be transferred securely into a data warehouse for further analysis.

2.2 Materials and Methods

2.2.1 Description and context of terms used:

1. **CEVM –** Centre for Evidence-based Veterinary Medicine, academic support and office of supervision for the research.
2. **Practice management software (PMS) system:** A computer based electronic health record management system, usually in the form of a database system supported by a commercial system provider.
3. **Electronic patient record (EPR):** The complete record for a patient within the PMS of a veterinary practice. The EPR should contain all signalment and clinical information, including historical data, recorded for each individual animal registered.
4. **Extracted EPRs:** For the purposes of this research an extracted EPR refers to a <u>collection</u> of patient records extracted from a PMS system (CEVM PMS or Sentinel PMS) using a predefined selection criteria.
5. **Patient history:** A complete health record for an <u>individual</u> patient created by the chosen veterinarian usually for the purpose of continued healthcare, inventory management and billing. Should contain all details recorded within the PMS for a single patient during their time registered at one veterinary practice.
6. **Animal ID:** A unique animal ID number is generated automatically by the PMS system and allocated to each patient whose details are entered into to the PMS EPR.
7. **Animal Record:** The 'Animal record' is the name of a table containing all individual patient records extracted and stored in the Data warehouse. The storage system uses the animal ID number extracted to create a unique record per patient of signalment details.
8. **Visit:** The details recorded into a patient EPR as part of a single consultation or visit.
9. **Visit Record:** The visit record is a table of individual patient records extracted and stored in the Data warehouse. The Data warehouse uses the Animal ID, Text Entry and Diagnosis fields extracted to create a unique record of visits per patient including the patient's clinical details which may include telephone call notes, test results, previous history, insurance claims or any other details added to the PMS EPR.
10. **Clinical notes or Clinical narrative:** The notes recorded in the PMS by the veterinarian, usually during the consultation, typed into the system manually and for the purpose of this study referred to as <u>free text.</u>
11. **Coded information:** The data added to a PMS or EPR using a numerical or textual code as shorthand (rather than free text) to describe and record a diagnosis, presenting

complaint or item for invoice. The code is usually selected from an integrated list created by the practice. For this research the codes were referred to as diagnostic code as this is the term chosen by the sentinel practice for their integrated list.

12. **Invoice:** The notes recorded in the PMS by the veterinarian, usually during the consultation, either typed into the system manually or added using a coded list containing details of all items to be charged. Any extracted invoice information did not contain any financial information.
13. **CEVM PMS system:** A pilot PMS system supplied to the CEVM by Vet-One (a web based or cloud Veterinary Management Software system provider) and created by the researcher (JJD) with (mock) patient records for the purpose of method development and extraction validation.
14. **Sentinel practice:** A real first opinion veterinary practice and collaborating partner to the Centre for Evidence-based Veterinary Medicine for the purpose of practice-based research.
15. **Sentinel PMS system:** The PMS system (also Vet One) of the sentinel practice used for the extraction research.
16. **Extensible Mark-up Language (XML) Schema**: A list of commands used to extract information from a database of records.
17. **Clinical Evidence (CE) XML Schema v1.05**: The schema designed for the extraction of EPRs from veterinary PMS systems for the purpose of this research and for inclusion into the Vet XML Consortium website.
18. **IE – Information Extraction:** A method of identifying and extracting information from an electronic resource or database. In this instance IE refers to the method used in this thesis and relates to extraction of selected patient information from a whole EPR.
19. **Data field:** Within a PMS or database each piece of information is recorded within a separate 'field'. The fields can be identified by a 'field heading' e.g. Practice ID or Animal ID. An XML schema is written to identify the information needed based on commands relating to field headings and field content, information for extraction was selected by requesting the data contained for each patient from selected fields.
20. **Data warehouse:** A relational database designed using Microsoft Access for the purpose of the secure storage of extracted patient information and aggregated electronic patient records.

The flow of work for Chapter 4 is presented in Figure 2-1 below,

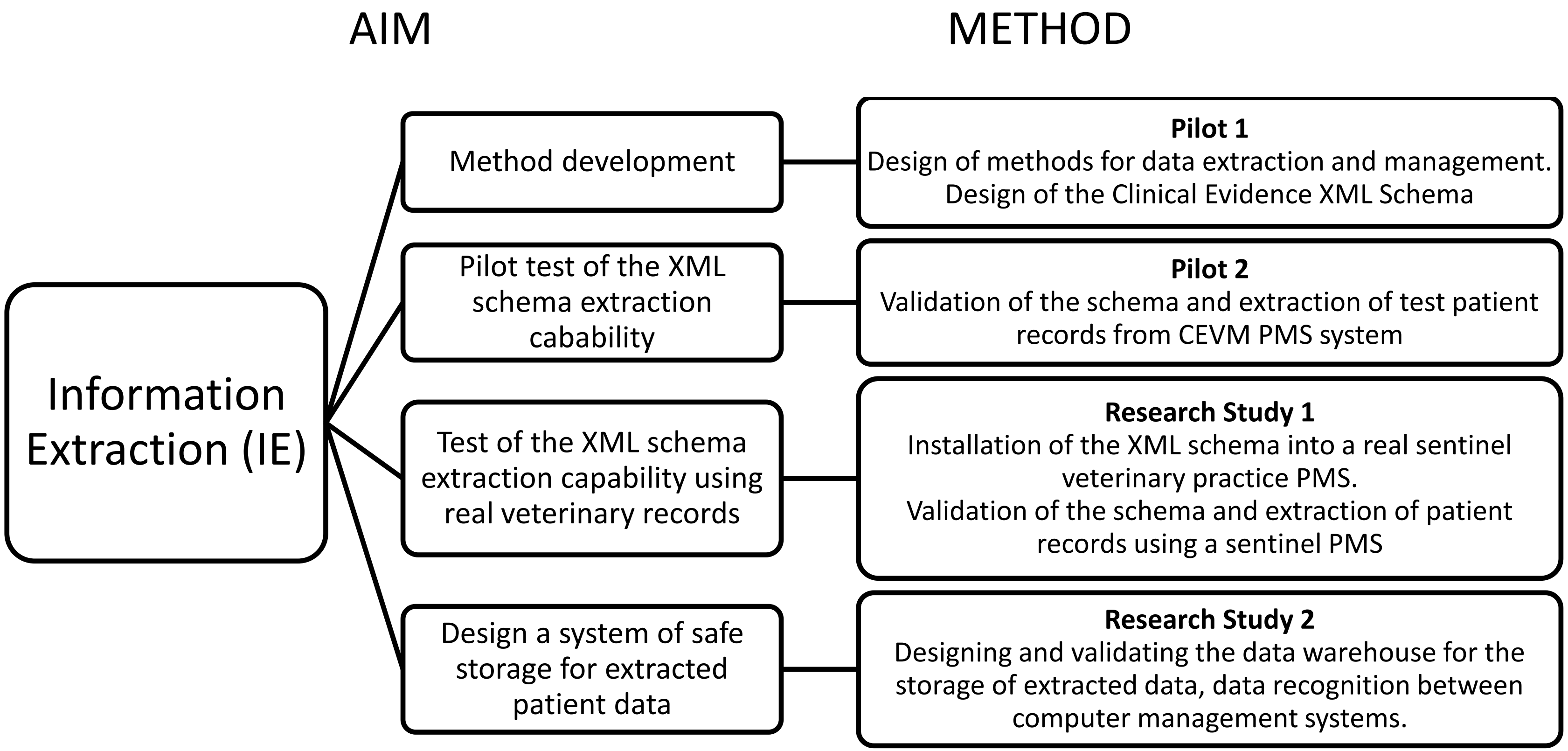


Figure 2-1 Flow of work for Chapter 4

2.2.2 Pilot 1 – Information Extraction Method Development

The aim of Pilot 1 was to identify an efficient and versatile system for data extraction which could be applied across many differing commercial PMS systems. The method of extraction would also be used as a tool to create a file of extracted data easily read by different database programs (Access, Excel) for the safe import and storage of extracted confidential patient records.

2.2.2.1 Extraction tool development

To extract all target data recorded in a typical EPR, an Extensible Mark-up Language (XML) Schema tool was chosen and associated extraction method designed. An XML Schema is essentially a list of instructions identifying the fields of interest, within the organisation of the data, recorded in a typical database or dataset (Quin, 2012).

Extensible Mark-up Language (XML) is a computer language that defines a set of rules for encoding documents in a format that is readable to both humans and computers. Extensible mark-up Language looks similar to Hypertext mark-ups Language (HTML) but XML is popular for data exchange because it can be used between systems and is simple and flexible, using common language and terminology (Katehakis et al., 2001, Quin, 2012).

An XML file of patient data, extracted from a commercial PMS system using a preliminary version of an XML schema, was supplied by a data management company called Vet Envoy[4] to pilot the use of an XML schema for the extraction method. The patient data file was provided to demonstrate how an XML schema could be applied to a veterinary PMS system to extract patient records and also to show the type of data which could be extracted. Vet Envoy also provided a file of extracted data to test how versatile XML data is for use between different computer systems.

2.2.3 Pilot 2 – Information Extraction Method Validation (CEVM PMS)

The aim of pilot 2 was to (i) design a XML schema for the extraction method, (ii) validate the XML schema for the extraction of patient records from a Vet-One CEVM PMS system and (iii) to validate the precision of the extraction method by comparing the data contained within the Vet-One PMS system to that extracted.

2.2.3.1 XML Schema design

A PMS system was requested from Vet-One Veterinary Management Software[5] to be used for Pilot 2 for use in validation of the XML schema for

---

[4] www.vetenvoy.com

[5] www.vet-one.com

EPR extraction. Information was provided by Vet-One regarding the architecture of the PMS system database and the fields of data available for extraction. An XML schema was then designed in collaboration with Vet-One. The fields chosen for extraction were selected from the full list of data fields available. All fields containing personal information, such as name and address of the owner and the name of the dog, were avoided to ensure the data extracted was anonymous. The 21 data fields identified for extraction from the Vet-One PMS included all animal demographic fields (e.g. species, age, breed), consultation information (e.g. condition history, examination findings, and consultation outcome) and any treatments given.

2.2.3.2 The data fields

The details of the 21 fields in the Vet-One PMS system chosen for extraction were as follows;

2.2.3.2.1 Animal fields

1. [Practice ID] – The vet practice ID to which the animal is registered. An individual practice specific number, normally the practice RCVS number, chosen by the practice. The Practice ID chosen by the sentinel practice was the Practice name and Phone number, this was entered by the system administrator when the system was set up.

2. [Animal ID] A unique integer that is auto incremented by the PMS system every time a new animal is added.
3. [Species] Animal species
4. [Breed] Animal breed
5. [Gender & Neuter status] Male entire, Female entire, Male neutered, Female neutered.
6. [Notable Conditions] A field for recording notes e.g. Allergy information.
7. [Remarks] A field for recording animal related notes e.g. may be aggressive.
8. [Deceased] An information field (Yes/No)
9. [Dangerous] An information field (Yes/No)
10. [Insured] An information field (Yes/No)
11. [Date of Birth] The date the animal was born if known.
12. [Body Weight] Animals last measured weight
13. [Body Weight Units] The unit used to record the body weight of the animal, usually Kilograms (Kg)
14. [Last Weight Date] The date the animal was last weighed
15. [Registration Date] The date the animal was registered with the practice and added to the PMS system.

2.2.3.2.2 Consultation fields

16. [Date] The date of the last entry to the patient's record.

17. [Time] The time of the last entry to the patient's record.
18. [Entered By ID] The numerical ID for the person who entered the data.
19. [Text Entry] A free text field where information can be typed in by hand usually used to enter details of the consultation or any information relating to the health of the animal. This field can also be used to add previous history, insurance details or test results.
20. [Diagnosis] A practice specific coded entry field for diagnosis or to record treatment or prescriptions to be invoiced. No costs or payment details were requested.
21. [Venom Code][6] A coded entry for diagnosis or clinical signs integrated into the system by the Venom coding group.

The XML schema was written by the Vet-One team with direction from the author and CEVM team. The design of the schema took 18 months to reach a final validated version and there were many revisions until the fifth and final version was created, Figure 2-2, termed Clinical Evidence (CE) XML Schema v1.0.5.

[6] Royal Veterinary College VetCompass and Venom Coding Group

```
- <xs:schema targetNamespace="http://www.vetxml.org/schemas/ClinicalData"
xmlns="http://www.vetxml.org/schemas/ClinicalData" xmlns:xs="http://www.w3.org/2001/XMLSchema">
- <xs:element name="ClinicalData">
- <xs:complexType>
- <xs:sequence>
- <xs:element name="Identification" minOccurs="0" maxOccurs="1">
- <xs:complexType>
- <xs:all>
  <xs:element name="PracticeID" type="xs:string" minOccurs="0" maxOccurs="1" />
  </xs:all>
  </xs:complexType>
  </xs:element>
- <xs:element name="ClinicalDataRecord" minOccurs="0" maxOccurs="unbounded">
- <xs:complexType>
- <xs:all>
- <xs:element name="AnimalDetails" minOccurs="0" maxOccurs="1">
- <xs:complexType>
- <xs:all>
  <xs:element name="AnimalID" type="xs:string" minOccurs="0" maxOccurs="1" />
  <xs:element name="Species" type="xs:string" minOccurs="0" maxOccurs="1" />
  <xs:element name="Breed" type="xs:string" minOccurs="0" maxOccurs="1" />
  <xs:element name="DateOfBirth" type="xs:date" minOccurs="0" maxOccurs="1" />
  <xs:element name="Gender" type="xs:string" minOccurs="0" maxOccurs="1" />
  <xs:element name="Deceased" type="xs:boolean" minOccurs="0" maxOccurs="1" />
  <xs:element name="Dangerous" type="xs:boolean" minOccurs="0" maxOccurs="1" />
  <xs:element name="Insured" type="xs:boolean" minOccurs="0" maxOccurs="1" />
  <xs:element name="NotableConditions" type="xs:string" minOccurs="0" maxOccurs="1" />
  <xs:element name="Remarks" type="xs:string" minOccurs="0" maxOccurs="1" />
  <xs:element name="LastWeight" type="xs:double" minOccurs="0" maxOccurs="1" />
  <xs:element name="LastWeightUnits" type="xs:string" minOccurs="0" maxOccurs="1" />
  <xs:element name="LastWeightDate" type="xs:date" minOccurs="0" maxOccurs="1" />
  <xs:element name="RegistrationDate" type="xs:date" minOccurs="0" maxOccurs="1" />
  </xs:all>
  </xs:complexType>
  </xs:element>
- <xs:element name="AnimalClinicalHistory" minOccurs="0" maxOccurs="1">
- <xs:complexType>
- <xs:sequence>
- <xs:element name="Entry" minOccurs="0" maxOccurs="unbounded">
- <xs:complexType>
- <xs:all>
  <xs:element name="DateTime" type="xs:dateTime" minOccurs="0" maxOccurs="1" />
  <xs:element name="EnteredByID" type="xs:string" minOccurs="0" maxOccurs="1" />
  <xs:element name="TextEntry" type="xs:string" minOccurs="0" maxOccurs="1" />
  <xs:element name="Diagnosis" type="xs:string" minOccurs="0" maxOccurs="1" />
  <xs:element name="VenomCode" type="xs:string" minOccurs="0" maxOccurs="1" />
  </xs:all>
  </xs:complexType>
  </xs:element>
  </xs:sequence>
  </xs:complexType>
  </xs:element>
  </xs:all>
  </xs:complexType>
  </xs:element>
  </xs:sequence>
  <xs:attribute name="version" type="xs:string" />
  </xs:complexType>
  </xs:element>
  </xs:schema>
```


Figure 2-2 Clinical Evidence XML schema final version v1.0.5

2.2.3.3 CEVM Vet-One PMS

A web based Vet-One PMS system Figure 2-3 was designed for pilot 2 (CEVM PMS) and access was provided via login ID for the researcher Julie Jones-Diette (JJD). Once the initial design of the XML schema (v1.0.0) was complete the schema was integrated into the CEVM PMS system architecture for validation.

An electronic patient record (EPR), containing a number of small animal records (n = 80 animals) was created by JJD and added to the CEVM PMS via the 'add new client' function on the PMS system main page, Figure 2-3. The details entered for each patient varied for species, age, breed, neuter status and clinical condition. A number of visit entries and associated invoices were created for each of the 80 patients. At least two visit entries were added to each of the 80 animals' records (to vary data entry, three of the 80 animal records, Animal ID 000045, 000008, 000009, had three visits added to their record with associated invoices). Visit details included some clinical history or presenting complaint and details of invoicing which contained treatments offered or prescribed. Invoicing detailed the items charged. The CEVM PMS therefore contained an EPR for 80 'mock' patients recorded as having at least 2 visits recorded with associated invoice records per patient. This produced a CEVM PMS EPR with a total of 326 separate record entries for extraction.

To replicate a real system extraction, the patient data entered into the CEVM PMS system was anonymised by omitting any signalment data such as pet name, owner name, address, and postcode or name of vet practice. Therefore no identifying information was recorded but the patients were all given an Animal ID number for reference. Invoice details were included to capture treatments offered or medications prescribed. However all details of billing or cost were removed. The veterinarian was recorded as JJD and all prescriptions added to the patients' records were selected by JJD from an integral catalogue created within the PMS system.

2.2.3.3.1 Installation of the XML Schema into the CEVM PMS

The schema was added to the CEVM PMS architecture by the Vet-One team. The schema commands were accessed via the 'reports' link on the main page of the system within the "Clinical Evidence data (VetXML)" link (see screen shot in Appendix 2. The Clinical Evidence XML schema was designed to use a dedicated timeframe for record extraction therefore required a start date and end date for data collection. The dates for extraction were 24th May 2011 to 1st June 2011 (9 days) which corresponded to the date of data entry for the 80 patient records.

The CEVM EPR extraction file was created by selecting the 'Generate Report' button (Appendix 2). This created a file of extracted EPRs which could be saved and imported to the data warehouse.

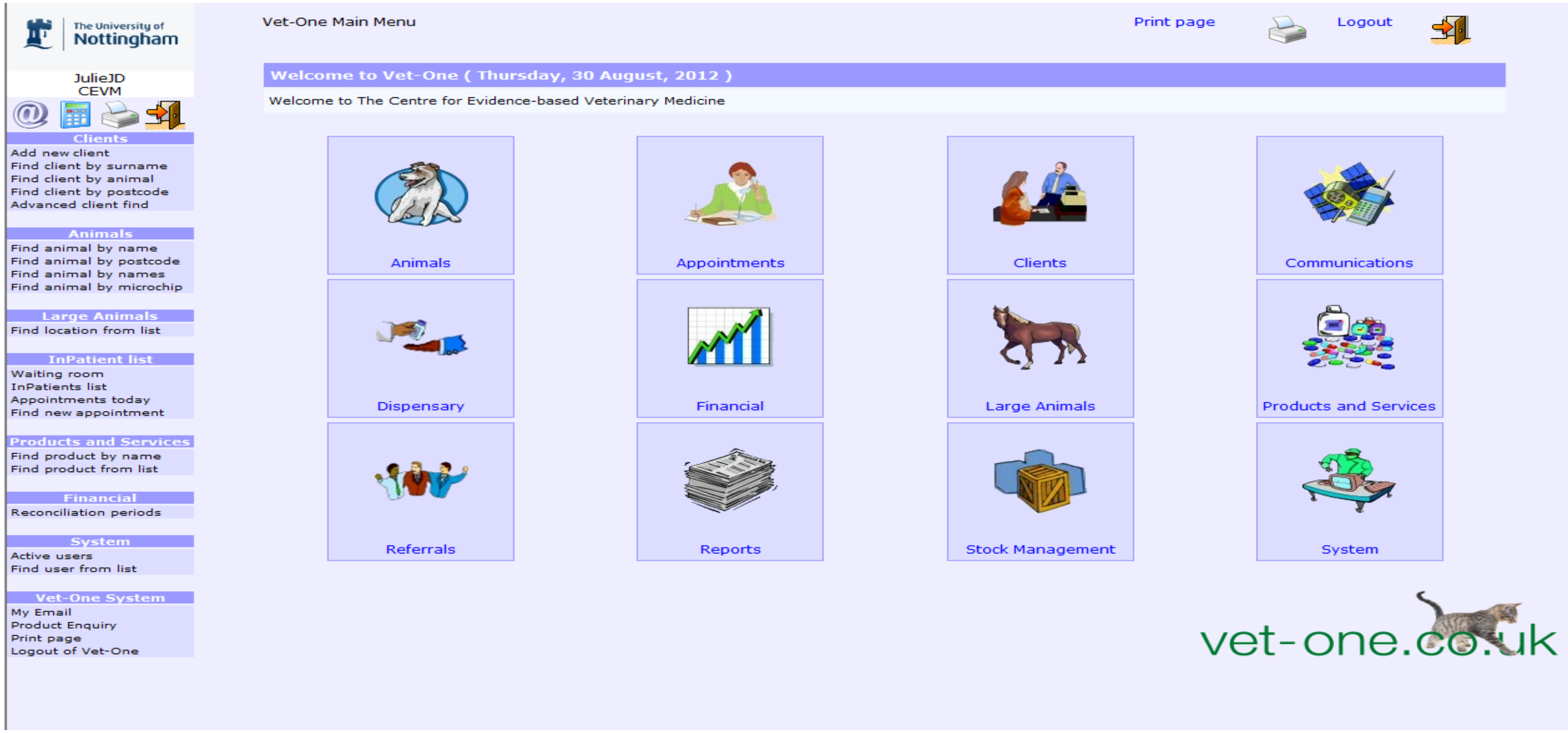


Figure 2-3 Screenshot of CEVM PMS system provided by Vet-One veterinary management software.

2.2.3.3.2 The opt-out function

As the extraction method was designed for future use on client owned patient data, an opt-out function was included in the study design to be integrated within the web based Vet-One PMS system. The opt-out function ensured any client wishing to be excluded from the research could do so on request. By selecting the opt-out field on a patient's record the schema would ignore the record when running the requested extraction for all future extractions. The opt-out process would not be reversed unless specifically requested by the client.

To validate the opt-out function, records within the CEVM PMS were selected and the opt-out field checked to exclude the record from the extraction. A report was generated to validate the operation of the opt-out function.

A further test of the system was created by adding new information to records where opt-out had been checked. This was to ensure the addition of new data to a record did not trigger their details for extraction in the future. A second report was generated to confirm that adding information did not lead to the extraction of data for those with the opt-out field checked.

2.2.3.3.3 Validation of the extraction method

Data from the CEVM PMS extraction were used to validate the functionality of the Clinical Evidence XML Schema and the suitability of the data warehouse for storage of the extracted data.

To ensure all information contained within the CEVM EPR was present in the extracted dataset, the information extracted from the CEVM EPRs was compared to the information printed out from the EPR of the PMS system (n=326 records). As each and every record was compared it was possible to confirm whether animal information, consultation information and invoicing information had been extracted completely for all records. The number of patient records present was compared to those extracted, the information for each field extracted was compared to those present in the paper print out and the animals that were opted out were checked to ensure their data did not appear in the record.

2.2.4 Research Study 1 – Information Extraction Method Validation (Sentinel PMS)

The aim of research study 1 was to validate a method to assimilate small animal patient data from a real first opinion small animal practice, including population data such as information on species and visit frequency. The schema developed was used as a tool to identify and extract all information required and the information extracted was used to create a data warehouse of patient records for population-based studies and to support practice-based research

2.2.4.1 Installation of the XML schema into a sentinel PMS

A working veterinary sentinel practice of the CEVM was recruited to enable the integration of the XML schema into a real practice PMS. The sentinel practice was a mixed animal first opinion practice in the south of England with three veterinarians working across two clinics. The practice agreed to provide access to their EPR for the purpose of extraction in a manner consistent with the ethical approval of the study. The XML schema was integrated by the PMS provider into the sentinel PMS and instructions for use were provided to the veterinarians at the practice by JJD (Appendix 3 and Appendix 4).

The extraction criterion was agreed to be Monday to Sunday and to include all clients examined at the practice within working hours (8-7pm). The file was sent by the senior veterinarian to JJD on the Monday morning of the week following data collection. The data collection periods ran in two 8 week blocks;

Data extraction period 1 January 16th 2012 – March 11th 2012

Data extraction period 2 May 2nd 2012 – June 25th 2012

At the end of each week during the data collection period an XML file of EPRs was extracted and forwarded to JJD by the senior veterinarian of the practice. The XML file was imported into an Excel spread sheet and cleaned for transfer to the data warehouse for storage.

After the data collection period was complete and initial analysis of the extraction had been performed, the CEVM research team and the sentinel practice veterinary team arranged a meeting to discuss the functionality of the method in a real working practice. The initial findings of the analysis were presented to the sentinel practice. A feedback session with the sentinel practice allowed an open discussion of how well received the process was by the clients and staff of the practice.

2.2.4.2 Informing clients of the study protocol

During data collection in the sentinel practice, posters were placed in each of the waiting rooms of the two clinics to inform all clients visiting the practice about the on-going project. The posters also informed clients that if they wished to opt out they could ask their vet to exclude the patient prior to or during the consultation. The consulting vet had instructions to select the opt-out function within the client's patient record for any clients wishing to be removed from the study (see Appendix 4 for instructions).

2.2.4.3 Exclusion criteria

As it was not possible to confirm from extracted data if an out of hours consultation had occurred in house or away from the practice, all clients examined out of hours were excluded from the data collection by the veterinarian who had instructions to use the opt-out function. This was to ensure all client information extracted was with informed consent. Finally only companion animals were included in the study, any farm animal consultations where the animal was not classed as a pet would be excluded from the study again by the vet selection to opt-out the patient.

2.2.4.4 Validation of the extraction method

To validate the practicality of the method for use in a real working practice, one to one feedback sessions were arranged to allow the sentinel practice veterinary team to comment on the system.

Once the sentinel PMS EPR extraction was complete for the first 8 week data collection period, a sample of records were chosen (10%; n = 252) to be printed out in full from the sentinel practice PMS. The printed records were used to compare the content against the extracted records to validate the precision of the XML schema for record extraction. Comparisons of basic data were performed including patient signalment information such as age sex species and breed and also the free text notes and any diagnosis or treatment information.

The completed data set contained a total of 2519 records. The randomisation function in Microsoft Excel was used to create a list of 252 random numbers from the total. These numbers were used to select a sample (10%, n = 252) of corresponding records for validation; records were selected from consultation information and invoice records allowing validation of both data sources. The sentinel practice provided paper records printed from the PMS for the random sample of 252 visit records to compare to the extracted visit record.

2.2.5 Research Study 2 – The Data Warehouse

Microsoft Office Access was the base program used to design the data warehouse for storage of the extracted data. All data extracted, from both the CEVM PMS and the sentinel PMS, were initially imported into Microsoft Excel where analysis and data cleaning was performed.

The data warehouse allowed separation of the data into two tables of information; (i) Animal table containing animal ID and further patient signalment data and (ii) Visit table containing animal ID linked to information collected during the consultation or invoice transactions. A relational hierarchy of tables was created within the data warehouse which allowed the data to be filed separated but linked for analysis with a "one animal ID with many visits" relationship. This relational separation ensured the data analysis identified each individual animal only once. However each animal may have had many visits to the vet during the data collection period therefore it was essential these visits were identifiable as isolated events but could still be linked back to the animal record.

A 'Visit ID' field was a **primary key** for the relationship between the database tables, which means the database recognised it as an individual record. A visit ID was added to the Visit table for each of the consultation records present.

Visit ID referred to a single consultation or patient record and was a unique number identified by the database as a separate entry which was essential for later content analysis and text mining.

The relationship for the separate tables was as follows;

Animal Table – *Query 1;* Practice ID, Animal ID, Species, Date of Birth, Gender, Deceased, Dangerous, Insured, Notable Conditions, Remarks, Registration Date. This table and query have Practice ID and Animal ID as the primary keys.

Visit Table – *Query 2;* Visit ID, Practice ID, Animal ID, Species, Breed, Last Weight, Last Weight Units, Last Weight Date, Date, Time, Entered By ID, Text Entry, Diagnosis, Venom Code. This table and query have Visit ID as the primary key.

The data fields selected for extraction have been presented previously with a brief explanation (4.2.2.2 **The data fields**). However their formatting within the field is important for explaining the development of the data warehouse fields and their importance for data organisation and analysis.

The 'Animal ID' field is vital for the separation of the patient data into the relational hierarchy within the database warehouse where the information is stored. The information extracted is anonymous to ensure it meets with client confidentiality and The Data Protection Act 1998. Therefore only the Animal ID was available to allow linkage to other data within the database for analysis.

The 'Date of Birth' field was extracted and recorded as dd/mm/yyyy. However it should be noted often date of birth may be entered as age therefore the default for the Vet-One PMS system date of birth standard format is the day the information is entered, followed by the month and year of birth calculated using the age.

The date the animal was registered [Registration Date] was a date field (dd/mm/yyyy). The Date and Time fields in Vet-One are formatted as a text field and are extracted as one field within the schema as follows; 2012-01-23T09:18:49.000Z. However this was separated after extraction into two fields [Date] & [Time] to allow analysis of the data by day.

The 'Entered By ID' is an individual number field allocated to each practicing vet, vet nurse or member of the veterinary team who are likely to input data. The ID number is auto incremented by the PMS system to ensure the numbers are unique within the practice. However ID details may be duplicated across practices of differing heritage that also have the Vet-One PMS system.

The ‘Text Entry’ field contains all the details recorded during the consultation by the veterinarian as free text, including any presenting complaint, clinical signs or recommended treatment. This field may also contain vaccination details including batch numbers of vaccines used, scanned in laboratory reports, previous history from the animals’ former veterinarian, referral reports and outcome data, insurance claims, details of any samples taken and any other important notes or details to be associated with the animals’ records.

The ‘Diagnosis’ field may include a practice specific coded entry for diagnosis or simply a code to identify the charging details including any invoice information (costs and payment details are not extracted). Finally the field entitled Venom Code was extracted which is a second coded entry field for diagnosis or clinical signs installed by the RVC VetCompass group[7] into veterinary PMS systems to monitor diagnostic and breed data.

The Access tables and initial queries were designed by the author with support from Tracy Hassall-Jones[8].

---

[7] http://www.rvc.ac.uk/vetcompass

[8] Microsoft Access consultant and trainer for The University of Nottingham

2.2.6 Data Security & Ethics

The study was approved by the ethics committee of The School of Veterinary Medicine and Science, The University of Nottingham.

All data was the responsibility of the Centre for Evidence-based Veterinary Medicine (CEVM). All patient data was extracted and stored in the strictest of confidence. No external parties had access to the data. The data was anonymised post extraction, including client and veterinary details, and for future examination and retention by the CEVM. No veterinary data was available to external parties.

2.3 Results

2.3.1 Pilot 1 – Information Extraction Method Development

2.3.1.1 The XML Schema tool

The XML file provided by Vet Envoy contained 13 patient records which included signalment and clinical data. The schema used to extract the pilot 1 patient data included field requests for the name, address and post code of the client or owner of the pet or patient which did not meet requirements of confidentiality for this study and therefore could not be used. The Vet Envoy schema extraction did however produce an XML file of data in an ideal format and the 13 records imported successfully and completely into both a Microsoft Excel database and Microsoft Access database.

2.3.2 Pilot 2 – Information Extraction Method Validation (CEVM PMS)

2.3.2.1 Validation of the EPR extraction from the CEVM PMS

All data was successfully extracted from the CEVM Vet-One system and were accepted and stored in full within the data warehouse. Eighty animals and 326 records were extracted in full. Validation of all extracted records using paper records for comparison found 100% accuracy. The invoice details and all treatments were also extracted in full. Validation confirmed no missing information (Table 2-1).

2.3.2.2 XML Schema design

2.3.2.2.1 The data fields for the CEVM PMS

The 21 data fields in the Vet-One PMS system architecture, chosen for the XML schema extraction tool, were validated for use by running the schema through the CEVM PMS system and checking no further information was required. All fields were extracted fully and all data within the fields were found to be accurate. The 21 fields selected for extraction were considered to be sufficient for the research.

Table 2-1 All animal records and corresponding extracted records for CEVM pilot test database extraction

Species	Sex	Neuter status	Number present	Number extracted
Cat	*M*	*E*	*4*	*4*
Cat	*F*	*E*	*4*	*4*
Cat	*M*	*N*	*20*	*20*
Cat	*F*	*N*	*11*	*11*
Cat	*-*	*N*	*1*	*1*
Cat	*-*	*-*	*1*	*1*
Total			**41**	**41**
Dog	*M*	*E*	*5*	*5*
Dog	*F*	*E*	*4*	*4*
Dog	*M*	*N*	*9*	*9*
Dog	*F*	*N*	*17*	*17*
Dog	*-*	*N*	*1*	*1*
Dog	*-*	*-*	*1*	*1*
Total			**37**	**37**
Bird	*-*	*-*	*2*	*2*
Total			**2**	**2**
Total No of animals			**80**	**80**

All data entered into CEVM Vet-One PMS system and corresponding extraction data.
M = Male, F = Female, N = Neutered, E = Entire

2.3.2.2.2 Validating the opt-out function

The opt-out function was tested by using sample records from the CEVM PMS system. Eleven patient records from the CEVM PMS system were altered to reflect they had opted out of the study (Table 2-2). All eleven patient records were successfully excluded from the extraction, regardless of additional data added to the record.

Table 2-2 All animals with opt-out indicated on their clinical record within the CEVM PMS system.

Owner ID	Animal ID	Details	Extracted
000001	1784533	Opted out, added clinical history, added invoice, closed record.	NO
000002	1784534	Checked record was opted in, added clinical history, added invoice, closed record.	YES
000003	1784536	Opted out, added venom code onto history, added invoice, closed record.	NO
000004	1784535	Added venom code onto history, added invoice, closed record.	YES
000005	1784537	Opted out, added history, added invoice, closed record.	NO
000006	1784538	Added to history, added invoice, closed record.	YES
000007	1784539	Opted out, added history, closed record.	NO
000008	1784540	Added diagnosis code and invoice to record then closed record.	YES
000009	1784541	Opted out, added notes, added invoice, closed record. Opted in, added more notes, closed record	NO
000009	1784541	Opted in, added more notes, closed record	YES
000010	1784542	Add notes, invoiced.	YES
000011	1784543	Opted out, added clinical history, added invoice, closed record.	NO
000011	1784543	Opened record, opted in, closed record.	YES

All data from CEVM Vet-One PMS system. Animal details and corresponding extraction data for validation of the opt-out option within the Vet-One PMS system.

2.3.3 Research Study 1 – Information Extraction Method Validation (Sentinel PMS)

2.3.3.1 Validation of EPR Extraction

Validation for extraction 'ease of use' in a working veterinary practice was achieved via a one to one feedback session with the veterinary team at the sentinel practice at the end of data collection.

The team were asked how they found the extraction method to use. They advised there were no issues with the system and it was easy to use. The feedback from the sentinel practice confirmed the system did not impact on the daily work routine. When asked if the work was of value to the practice they were very supportive of the method and the collection of data from their practice and were keen to continue to be involved and found the research findings very interesting and informative.

2.3.3.2 Validation of the XML Schema precision

The extracted data matched the information present on the paper records printed from the sentinel practice perfectly and the extraction was found to have 100% accuracy by direct visual comparison. The fields of information were an exact match as were the data contained within each data field. All free text was extracted in full as was the invoicing information, prescriptions and all recorded treatment given to the patient.

2.3.3.3 Data extracted from the Sentinel PMS

Extraction of data for Period 1 and 2 produced a combined total of 4946 extracted records. These records were composed of 2246 visit notes and 2700 invoices and involved 1279 individual animals.

The 2246 visit records included notes recorded as consultation (n=1858), test results/lab reports (n=292), insurance details (n=64), previous history (n=22), referral notes (n=6) and follow up appointments (n=4). Of the 1858 notes recorded as a consultation, 1624 were recorded simply as ‘consultation’, 83 as ‘vaccination’, and 14 as ‘phone calls’ or ‘phone consultation’, ‘examinations’ where a diagnostic code was recorded accounted for 137 records. A VeNom (veterinary nomenclature coding group) code was used for only 1% of visit records (n=30) and each venom code was combined with a practice specific diagnostic code.

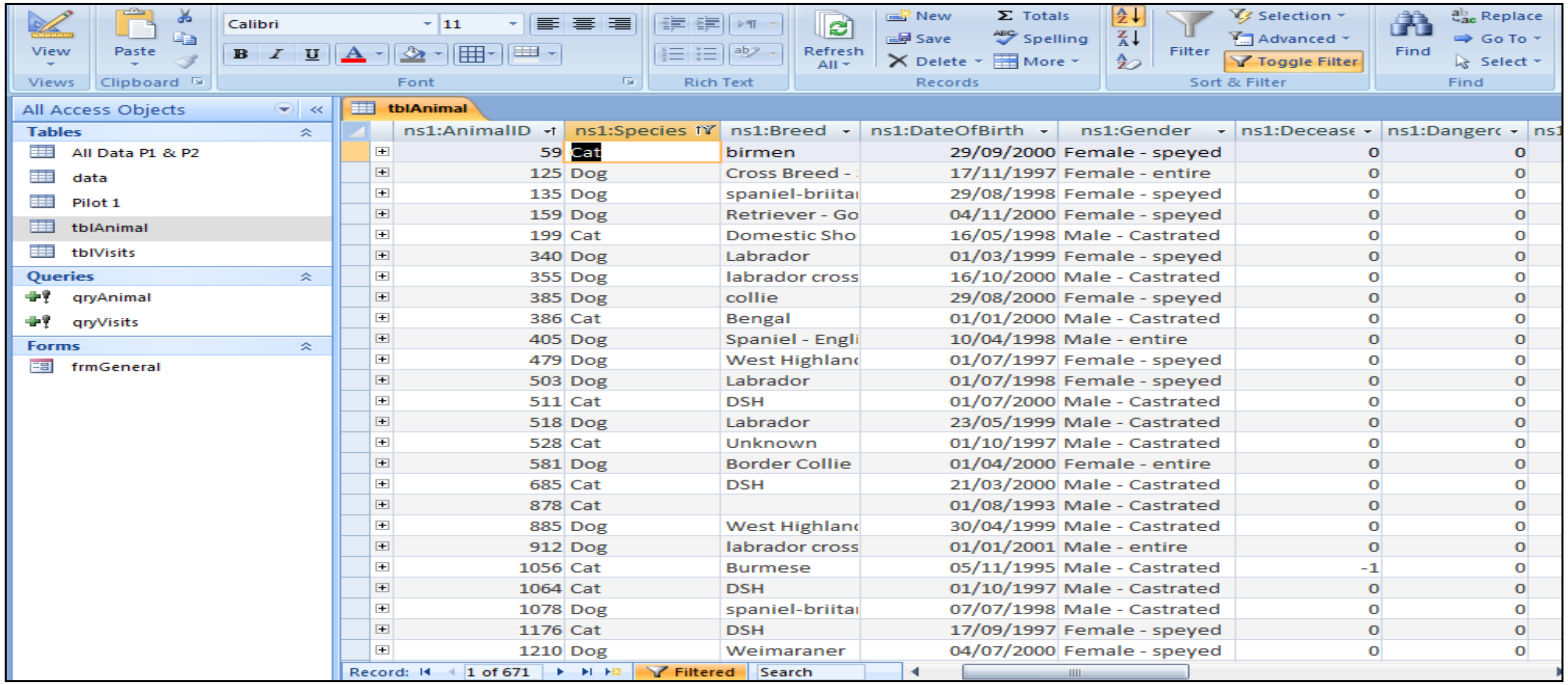
2.3.3.3.1 Animal Details

All animal records were extracted in a form which could be stored and managed within the data Warehouse (Figure 2-4). Period 1 (Jan to March 2012) extracted 2519 patient records in total, visit records and invoice information combined, from 775 animals. Period 2 (May to June 2012) extracted 2427 patient records in total, visit records and invoice information, from 822 animals.

This produced a dataset of combined animal records 1597 animals. The relational database separated out the individual animals examined from the combined total of 1597, as some animals visited the practice during both collection periods. Therefore the exact number of individual animals examined over both data collection periods combined was 1279 animals, Table 2-3. The animal species distribution included a number of farm animals n=114 (9%) birds, including poultry n=36 (3%) and blank entries n=45 (4%).

Table 2-3 Animal details for period 1 and period 2 data extraction from the sentinel practice PMS

Species	Period 1 n	Proportion (%)	Period 2 n	Proportion (%)	Period 1 & 2 N	Proportion (%)
Dogs	439	0.57	469	0.57	693	0.54
Cats	198	0.25	217	0.26	336	0.26
Rabbits	25	0.03	16	0.02	38	0.03
Other	113	0.15	120	0.15	212	0.17
Total	**775**		**822**		**1279**	



ns1:AnimalID	ns1:Species	ns1:Breed	ns1:DateOfBirth	ns1:Gender	ns1:Decease	ns1:Dangerc
59	Cat	birmen	29/09/2000	Female - speyed	0	0
125	Dog	Cross Breed -	17/11/1997	Female - entire	0	0
135	Dog	spaniel-briitai	29/08/1998	Female - speyed	0	0
159	Dog	Retriever - Go	04/11/2000	Female - speyed	0	0
199	Cat	Domestic Sho	16/05/1998	Male - Castrated	0	0
340	Dog	Labrador	01/03/1999	Female - speyed	0	0
355	Dog	labrador cross	16/10/2000	Male - Castrated	0	0
385	Dog	collie	29/08/2000	Female - speyed	0	0
386	Cat	Bengal	01/01/2000	Male - Castrated	0	0
405	Dog	Spaniel - Engli	10/04/1998	Male - entire	0	0
479	Dog	West Highlan	01/07/1997	Female - speyed	0	0
503	Dog	Labrador	01/07/1998	Female - speyed	0	0
511	Cat	DSH	01/07/2000	Male - Castrated	0	0
518	Dog	Labrador	23/05/1999	Male - Castrated	0	0
528	Cat	Unknown	01/10/1997	Male - Castrated	0	0
581	Dog	Border Collie	01/04/2000	Female - entire	0	0
685	Cat	DSH	21/03/2000	Male - Castrated	0	0
878	Cat		01/08/1993	Male - Castrated	0	0
885	Dog	West Highlan	30/04/1999	Male - Castrated	0	0
912	Dog	labrador cross	01/01/2001	Male - entire	0	0
1056	Cat	Burmese	05/11/1995	Male - Castrated	-1	0
1064	Cat	DSH	01/10/1997	Male - Castrated	0	0
1078	Dog	spaniel-briitai	07/07/1998	Male - Castrated	0	0
1176	Cat	DSH	17/09/1997	Female - speyed	0	0
1210	Dog	Weimaraner	04/07/2000	Female - speyed	0	0

Figure 2-4 Patient details extracted from a sentinel PMS and stored in the CEVM data warehouse.

One hundred and twenty two cats made a total of 299 visits in period 1. Fifty nine percent of cats only visited the veterinary practice once (Figure 2-5). However 21% visited on two occasions, 12% on 3 occasions and 5% visited on four occasions up to a maximum of 10 visits in 8 weeks.

Three hundred and twenty dogs made a total of 683 visits to the veterinary practice during period 1. Fifty two percent visited the veterinary practice on only one occasion, 20% on two occasions, 13% on three occasions and 5% on four occasions. A maximum of 15 repeat visits for one dog was recorded however this because repeated chemotherapy treatment was required (Figure 2-5).

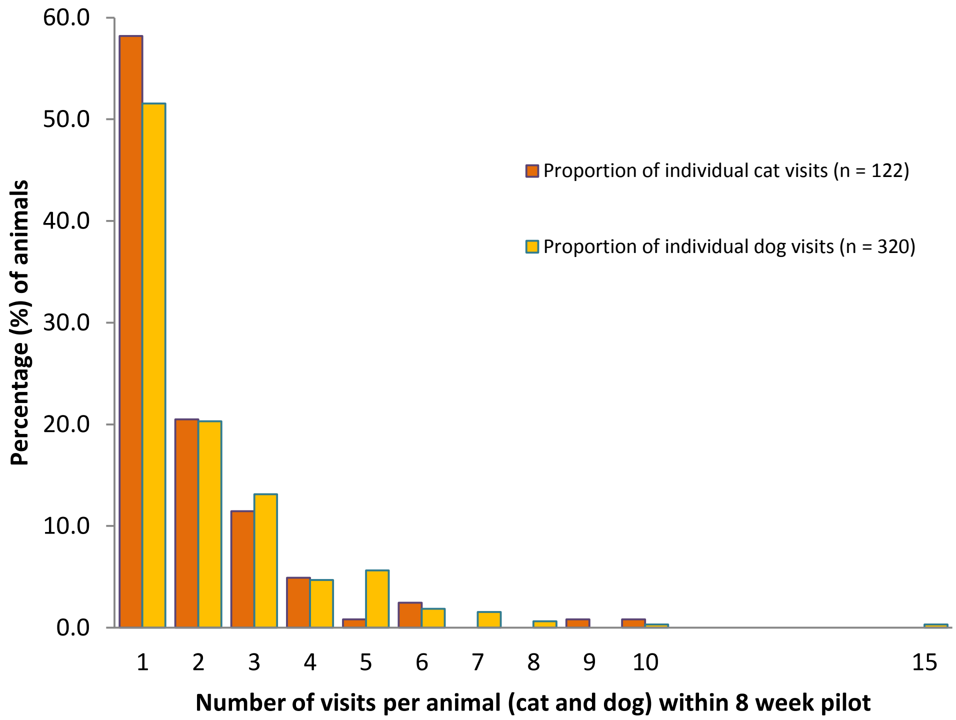


Figure 2-5 Visit frequency analysis for data extracted from a sentinel practice over 8weeks of data collection, cats n=122 dog n=320 (min visits 1, max visits = 15)

2.3.3.3.2 The Text Entry field

The extracted 'Text Entry' field of the record (n = 4946 extracted entries) containing the free text was 100% precise when compared to the information on the 10% sample of printed records from the sentinel practice PMS, Figure 2-6. The Text entry field mostly contained typed notes (free text) recorded during consultation (n=1862; 37.6%). The field also contained documents added to the record such as lab reports, previous histories, insurance details and insurance notes (n=384; 7.8%).

Just over half of the records were coded as an invoice (n = 2700; 55%). The details recorded in the text entry of an invoice provided information on treatments or drugs prescribed or services offered, e.g. surgery, by the veterinarian during the consultation. As requested, no billing information or financial details were included in the extracted data.

2.3.3.3.3 Date and Time of Consultation

The date and time for each data entry was automatically recorded by the PMS and extracted for each field of data and imported successfully as a single field into the data warehouse.

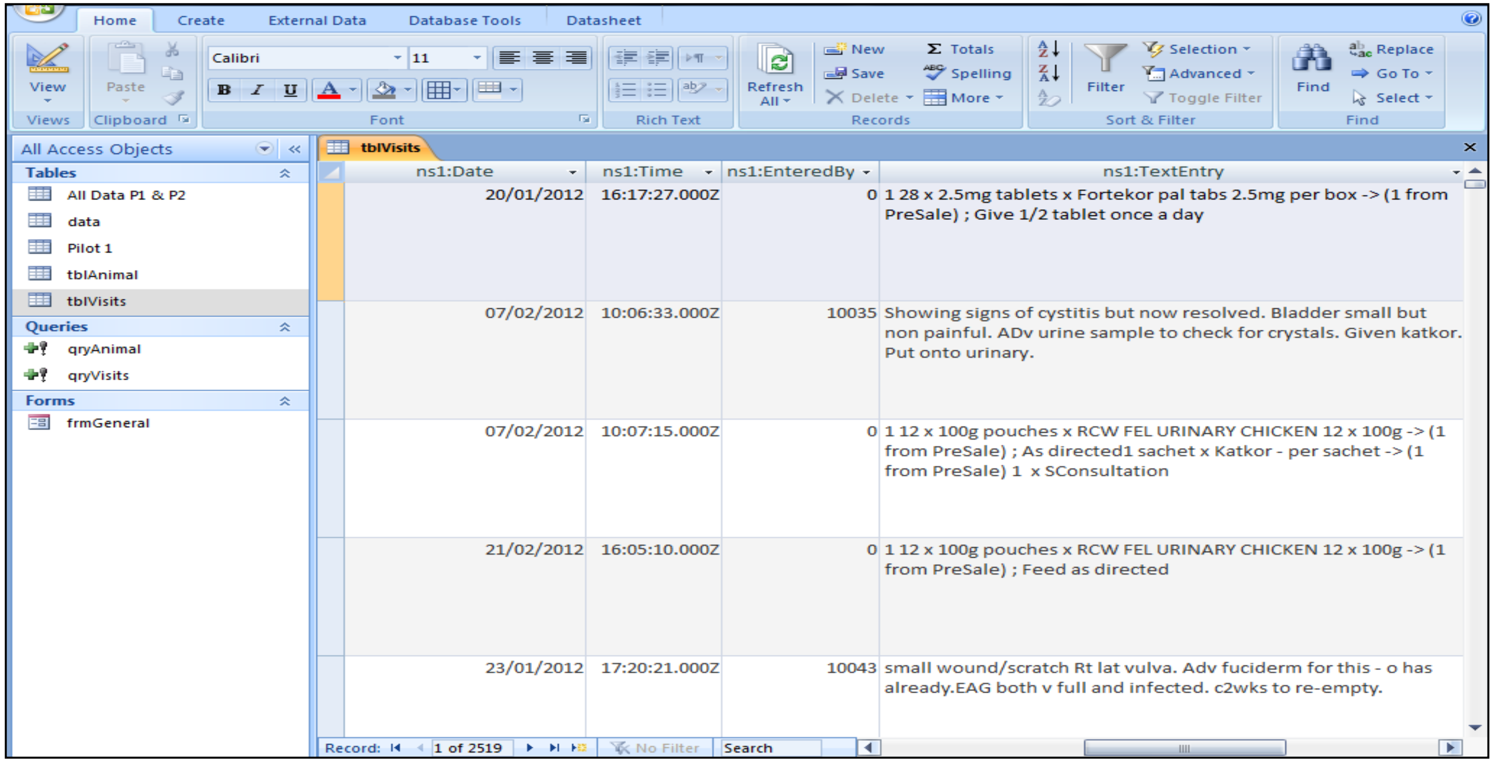


Figure 2-6 Screenshot of extracted clinical free text from the sentinel PMS system

2.3.3.3.4 Diagnosis Coded Entry

The diagnosis coded entry was extracted with 100% precision when compared to the random sample of paper records (n=252). The Practice diagnosis code list and venom lists were both extracted fully.

The used of the diagnostic codes varied between veterinarians. Diagnostic practice codes were recorded for 137/2246 examination records excluding invoice records. Only 6% of records, for Period 1 and Period 2 combined, contained a diagnostic code and on 83% of the occasions when a code was used it was by only one veterinarian. The system extracted 30/2246 venom codes, indicating a 1.3% use, again by the same single veterinarian.

2.3.3.3.5 Practice ID

The practice ID was extracted and stored as a primary key within the data warehouse allowing for unique ID for the practice and associated data.

2.3.3.3.6 Entered By ID

Three veterinarians were involved in data entry at the sentinel practice with additional members of the practice team responsible for collecting data during the two 8 week periods including veterinarians, nursing staff and administrators. In Period 1, 14 members of staff collected data; in Period 2, 11 members of staff were involved in data collection. Overall there was a total of 15 Entered By IDs recorded with 10 people involved in both data collection periods (Table 2-4).

Table 2-4 List of 'Entered By ID' for members of the veterinary team at the sentinel practice responsible for data input per collection period.

Number of staff with access to PMS	Entered By ID	Data Collection period 1 (Wk. 1-8)	Data Collection period 2 (Wk. 1-8)
1	Blank	Blank	Blank
2	0	0	0
3	2	2	2
4	13	13	
5	10035	10035	10035
6	10037	10037	
7	10039	10039	
8	10041	10041	10041
9	10043	10043	10043
10	10045	10045	10045
11	10047	10047	10047
12	10048	10048	
13	10050	10050	10050
14	10051	10051	10051
15	10052	10052	10052
16	**10057**		**10057**

An 'Entered By ID' is allocated by the team member at the practice with input rights for the PMS updates. ID numbers are allocated to all members of the practice who may be required to add data to the computer records. This may include veterinarians, nurses, receptionists, practice managers and admin teams.

2.3.4 Research Study 2 - The Data Warehouse

The files extracted from the CEVM PMS system were imported into an Excel spread sheet. They were cleaned and then imported into the Access Data warehouse. The data management systems performed well and the data warehouse successfully imported all requested fields. However the free text was found to be incomplete initially. The Access data warehouse restricted the free text import due to the 255 character limit of the text field and additional formatting of the original text. The import to Access was resistant to a change from a text to a memo file, which would normally allow for greater text content. Converting the text field to a memo prior to import did not help either as extra character script (see Appendix 5) within the extracted text appeared to cause problems for Access in character recognition.

In the PMS system some records were found to be many pages long particularly where reports or previous history has been pasted into the record and this also appeared to increase the script present in the extracted text. Fortunately all data was imported in full into the Excel database and could be cleaned within Excel before importing to the data warehouse in Access. Cleaning included removal of all character script and altering the background of the free text field to a text field allowed the transfer of the data.

2.4 Discussion

The work presented here suggests that using an XML schema for the extraction of clinical information from veterinary practice PMS could provide the versatile approach needed for combining data from many different veterinary practices utilising many different PMS systems. The extraction methodology produced excellent results with 100% precision and once integrated into the PMS system required very little involvement from the veterinarians at the practice. Finally, the transfer of extracted data into a suitable storage system for further analysis was also successfully achieved using a simple Access relational database and an Excel import function for the XML files.

2.4.1 XML Schema

An XML Schema was chosen for the project due to its flexibility to work across different computer systems and was preferred for this research over other methods currently in use; such as the introduction of a clinical coding system (Summers et al., 2010, O'Neill et al., 2012), bespoke data collection software with query analysis (Lund et al., 1999, Faunt et al., 2007, Hippisley-Cox and Stables, 2011, Anholt et al., 2012, Anholt, 2013, Robinson et al., 2008) or using an embedded questionnaire (Radford et al., 2010).

The benefit of an XML schema over these other processes is not only the flexibility to work across many systems simultaneously but also that there is very little impact on the practitioners work load as the system is extracting what is already there rather than requesting any new or additional information. In recent years XML data transfer has increased in value as a tool for the exchange of information (Katehakis et al., 2001) particularly in the management of healthcare information (Patterson et al., 2002). The XML schema in this study worked perfectly as a tool to select and extract clinical data from the main database within the veterinary practice. The schema was found to have an excellent level of precision when extracted data was compared to the EPR, suggesting the method can be used with confidence.

The design of the XML schema was a lengthy process however now designed the process of modification for integration into other PMS systems is likely to be a much quicker process. The penultimate version of the schema (V1.0.4) worked perfectly, however certain changes were made to data field extraction for reasons of confidentiality, to ensure collection of a complete EPR and also to include any practice specific coded information. These modifications were made very quickly once the schema was designed. As a member of the Vet XML consortium[9], the Centre for Evidence-based

[9] Vet XML Consortium, Central Ltd, Elmtree Business Park, Elmswell, Bury St Edmunds, IP30 9HR

Veterinary Medicine (CEVM) was able to collaborate with a PMS system design team, interested in data sharing, to develop a method using XML language for data extraction. However, although the method tested here has potential to be integrated in its current format into many different PMS systems this is still to be tested and would be a recommendation for future work. It would also be advisable to investigate how much impact this would have on the PMS system provider and PMS systems function. The study presented here found no effect of the schema on the working of the PMS system. However future work will need to investigate if this is the case for other system types.

2.4.2 The Vet-One PMS system

The Vet-One PMS system installed for the CEVM to pilot the extraction was essential to the project aims. Its inclusion allowed validation of the schema and manipulation of the healthcare data within the database, ensuring security measures were working, such as the opt-out tool, before the method was integrated into a working sentinel practice.

2.4.3 Pilot 1 – Information Extraction Method Development

The initial findings from Pilot 1 allowed confidence that XML schema extraction tool would be useful to develop further. The 13 records provided by Vet Envoy allowed a study of a simple XML schema as a tool for extraction and also allowed a test of data import into a database for storage. The data files sent by Vet Envoy, for initial validation of the extraction method, provided a useful learning experience particularly with respect to the design of the database for the data warehouse. The files also allowed simple and efficient familiarity with XML files and import functionality.

2.4.4 Pilot 2 – Information Extraction Method Validation (CEVM PMS)

The PMS system provided by Vet-One was an excellent tool and allowed the validation of the extraction method and the data warehouse. The strength of this pilot study lies in the fact that as the data within the PMS system was added by the researcher, validation was efficient and accurate as all content was known and recorded prior to extraction for comparison. This allowed a superior level of confidence in the precision of the extraction for the CEVM system and for the research study.

2.4.5 Research Study 1 – Information Extraction Method Validation (Sentinel PMS)

The feedback session with the sentinel practice allowed an open discussion from the point of view of the practice. The comments received were very valuable for method development but also to understand the impact, if any, the method would have on the day to day workings of the practice and their clients. The feedback received from the sentinel practice was very positive.

It is encouraging that no clients chose to opt-out of the study from the sentinel practice. The findings of another study performed by members of the

CEVM (*unpublished data,* N. Robinson *Ph.D Thesis 2014*) also found a very low level of withdrawal from practice-based research generally, both by participating veterinarians or their clients.

Out of hours clients were not included in the study as it was assumed out of hours clients may not have opportunity to read the literature available in the waiting room during their appointment, particularly if the appointment was an emergency or if the veterinarian was required to make a home visit. Therefore it was the responsibility of the veterinarian to exclude these clients from the extraction by selecting the opt-out function. Although it is assumed the veterinarian excluded the data from the extraction, the timing of some consultations suggested they could be out of hours. As there was no way to confirm this within the limits of the current study, this would need to be a consideration for the future research design.

For all clients, the responsibility to opt-out of this research would be with the client or veterinarian which may not always be a reliable assumption, particularly when the vet is busy and research may not be their first priority. Therefore as a secondary safe guard for the ethics and confidentiality of the project no data was included in the analysis which did not meet the selection criteria and would be excluded prior to data analysis during the process of data cleaning.

Although there were a number of farm animals included with the extraction from the sentinel practice during the primary study it is not possible to confirm if the animals were seen at the practice. Although not essentially a problem with methodology for extraction of data, this does make data cleaning more difficult and time consuming. If the protocol is not followed more time would be required for checking and cleaning the information extracted. This highlights a limitation of the method as the researcher has very little control over what is recorded by the veterinarian. As such there may be a need for the researcher to work more closely with the veterinary team both with regard to training and on-going data collection.

The diagnostic codes added to the system by the sentinel practice for their own personal use appeared to have limited value. Only a few of the codes were used, with the veterinarians choosing general terms more often such as 'consultation'. More frequently, codes were not used at all, or occasionally when the coded descriptions were used they were recorded inconsistently between veterinarians and often terms were duplicated due to synonymous use or simple differences in spelling.

The research therefore found very few consultations included a diagnostic code (6% of visit records) or a venom code (1% of visit records). In addition the codes were mostly used by only one veterinarian and each VeNom code was combined with a practice specific diagnostic code making an unnecessary duplication of information.

There is currently little available data on the use of the venom codes in the UK since their introduction, although their use has been discussed for some studies (Summers et al., 2009, Kearsley-Fleet et al., 2013). Therefore it is difficult to speculate why in our experience codes were not used by all veterinarians all of the time at the sentinel practice. However when discussing this with the veterinarians at the sentinel practice it was explained the practice specific codes had been created by one veterinarian to assist with their own data input and was not a common practice or requirement for all veterinarians to record codes at the sentinel practice. As use of the practice and venom codes was optional, this lack of agreed standard terminology within the practice may explain their limited use with the veterinarians maybe finding it difficult to find the code which best describes the animals condition or consultation outcome, choosing instead to use terms they are most familiar with to describe their findings. It is also difficult to code a consultation if there is no definitive diagnosis at the time of the animal's examination and research suggests often a definitive diagnosis is not immediate but discovered at a later time. Finally one would assume there has to be greater value to the practice veterinarians for using a diagnosis code, when a diagnosis is reached, rather than free text for there to be a change in recording behaviour, and this may not currently be the case.

Many individuals within a practice may have a responsibility to update the EPR throughout the day. Overall for the research study presented here there was a total of 15 different 'Entered By ID's' recorded. This multiple entry of data is unavoidable in a real working situation, particularly in a large practice, and is the nature of practice-based research. However it is an important consideration when examining the data extracted. It may be that the data is examined to understand the relationship between the veterinarian, the patient and the PMS system. If support-staff are also recording information, one must be aware of this and take care when drawing conclusions on ways of working based on just the data entered, as some entries may not be by veterinarians.

Additional issues with the extracted data included missing data. One example of this was a blank 'Species' field entry on a number of extracted EPRs. During the feedback session with the veterinary team they advised it was possible the animal's species was recorded in the 'animal name' field instead. The XML schema does not extract animal name as all data needs to be anonymous, therefore the data was missing from the extracted EPR. Knowledge of the ways of working within the practice and how the veterinary team enter data is very important for data extraction and data quality.

2.4.6 Research Study 2 – The Data Warehouse

The Data warehouse performed perfectly. The data once cleaned could be imported directly into the Access database and the queries within the system separated and filed the data into the tables as it was imported. All animals appeared only once within the Animal table and all data entries appeared in full within the Visit table, linked by animal to allow the relationship of one animal to many visits to be identified as required. The separation and organisation of information within the warehouse made the analysis and the calculation of basic data more simple and efficient.

The need to be consistent for the data to be imported and stored correctly within the data warehouse was however very apparent. Any changes made to the fields in the PMS would disrupt the data warehouse filing system. The primary fields selected for data separation and recombination within the Access database are essential, an example of this is the practice ID which is specific to each practice. The data warehouse is designed to use the Practice ID field to protect against mixing of practice data. Changes occurring within the PMS system Practice ID field during data extraction provided a challenge for data management. Any changes made by the practice to their practice ID mid-trial, which did occur on one week of the extraction in error, meant the data warehouse would not recognise the data and assumed a new practice had been added to the system creating a new dataset instead of combining the data with the previous set.

These issues emphasis the need for consistency during data collection and also the importance of communication between the researcher and the sentinel practice. Changes made to the PMS systems in veterinary practice may disrupt data collection and storage but which essentially are at the discretion of the PMS system providers and their clients. The limits of the method must therefore be understood and rely heavily on communication and the compliance of the sentinel practice and the PMS system provider.

2.4.7 Strengths, Limitations and Future work

The value of the data extracted will be the focus of the following chapters. The data extracted can only be as good as that originally recorded. Data extracted from veterinary practice EPRs is unlikely to include standardised data as there is currently no standardised recording of information in veterinary medicine. Instead there may be spelling mistakes, shorthand, nuances and gaps in information which is to be expected from a busy working practice.

Criticism has been made of the use of data recorded in EPRs for practice-based research in the human healthcare field because it may be incomplete (Harman et al., 2012, Bernstein et al., 1993) or because its completeness has not been validated (Bernstein et al., 1993, Stein et al., 2000, Hassey et al., 2001, Harman et al., 2012). Additionally there has been concern over data sharing and confidentiality (McCurdy, 2001) which may affect participation by the veterinary profession. Professional issues may also need to be addressed such as the concern over anonymity and confidential ways of working for both clients and vets[10], including concern over regulation of prescribing and the 'cascade'[11], legislation of treatment[12] & disclosed veterinary identity.

Although there are many obstacles to overcome, this method has been shown to be a highly successful with great value to veterinary epidemiology and practice-based research. The market share of Vet-One PMS in the UK is 3% which is approximately 70 veterinary practices. This offers a large data set for veterinary epidemiological studies and a great deal of data for practice-based research. As such the development and growth of the data warehouse and extraction operability would need to be considerable. Recommendations would include a web based system with enhanced data storage capability and integration of the XML schema into other sentinel practices for further method development. Closer control over data input, training for the

---

[10] http://www.rcvs.org.uk/advice-and-guidance/code-of-professional-conduct-for-veterinary-surgeons/#clients

[11] "The Cascade is a difficult concept to implement and enforce but we try to get the right balance between animal welfare, not interfering with the clinical judgement of the veterinary surgeon and ensuring that UK authorised veterinary medicinal products are used as the default where appropriate...........We contend that effective regulation provides a significant level of assurance to veterinary surgeons by ensuring veterinary medicines can be relied upon to do what it says on the label. This can only be a secure assumption if underpinned by adequate legislation and good science." Taken from www.vmd.defra.gov.uk/mswd/cascade

[12] http://www.noahcompendium.co.uk/Compendium/Overview/-45043.html

veterinary team and the design of the schema would need to be managed as any changes made by either the PMS provider or the veterinarian could affect the success of the method and value of the data collected.

2.5 Conclusion

The extraction method and storage within the data warehouse of the extracted data has yielded a great deal of information for analysis and practice-based research. The potential value to the veterinary profession and the opportunity for research is sizeable. The results of this group of studies, the success of the methodology and the high level of precision for the extraction system, provide great encouragement for the future of practice-based research utilising medical informatics and XML language technology.

Chapter Three

3 Text-mining and Content Analysis

3.1 Introduction

Information held within the electronic patient record (EPR), whether medical or veterinary, is a unique source of data for research (Hersh, 2003b). Veterinary patient information is generated from an interaction between an animal, an owner and a veterinarian and is recorded to document this encounter. There are other uses for this type of information; decision support, retrospective outcome research and quality assurance (Hersh, 2003b). The aggregation of patient data is also of great value to veterinary practice-based research to examine disease prevalence and incidence within the vet visiting pet population (Faunt et al., 2007).

3.1.1 Information Extraction

The automated system of retrieving data or information held in computers is often termed **information extraction (IE)** (Hersh, 2003b) and is most commonly mentioned in reference to the extraction of individual facts or small details from large collections of unstructured texts such as EPRs. IE is used most often for informatics research (the use of computer systems to support research) (Hersh, 2003b) or medical informatics (the use of computer systems to support medical or clinical research) (Meystre et al., 2008, Hersh, 2003b).

When the goal is more likely to be the retrieval of specific factual information such as whether a patient had a particular clinical sign or diagnosis, the requirement for accuracy is much higher. While the consequences of an incorrectly spelled word in the record of a patient may be modest for a GP or veterinarian, the consequences of incorrect information extraction from a clinical narrative, the free-text portion of the record, for clinical epidemiology or research can be much more serious. This could lead to an inappropriate recommendation in the care of a patient or an incorrect assessment of the efficacy of a treatment in a population (Hersh, 2003b, Shortcliffe and Blois, 2003, Anholt, 2013). One problem is that clinical narratives are usually written or dictated quickly in a shorthand style with misspellings and grammatical incompleteness. As a result of these problems Hripcsak et al., (1995) noted that vital information is usually "locked" in the clinical narrative as it is difficult to predict and identify using informatics techniques.

The alternative is the extraction and analysis of coded information within the EPR. Unfortunately the findings of the previous chapter would suggest that using a coding system to analyse veterinary clinical data has limited value. As the codes need to be structured, consistent and standardised across practice to be of use to practice-based research, the current lack of an agreed standardised veterinary clinical coding system within veterinary practice means this is not a viable option. In addition, Jollis et al. (1993) considered options for outcomes research using the human EPR and states there is good reason to attempt to analyse the free-text information, as the coded information "does not capture the richness or complexities of the patient and the course of their disease". In addition some research has shown disagreement between the data in coded and free-text portions of the human medical record (Stein et al., 2000).

Content Analysis (CA) is an informatics technique employed for the analysis of documents of textual information and usually describes an automated search and or processing of the content of an extracted record (Lam et al., 2007a); in the case of the following chapter an electronic patient record (EPR) extracted from a veterinary practice management system (PMS). Content analysis is often combined with **Text Mining (TM),** where text mining explains the process of extracting knowledge from the analysis of unstructured text (Shortcliffe and Blois, 2003, Hersh, 2003b, Chen et al., 2005, Meystre et al., 2008).

WordStat is a computational linguistics program which was used for this research as an automated content analysis system to test its suitability for the identification and separation of words or phrases from within the free text portion of pilot and sentinel EPRs. This software provides large scale content analysis in a very short time period and far exceeds the ability of human search and retrieve capability within the same timeframe (Lam et al., 2007a).

This chapter describes the establishment and validation of a method of content analysis applied to electronic patient records extracted from a veterinary PMS system.

The aims of the study were;

(i) To pilot and validate the use of a content analysis program (Wordstat software) to text mine the EPR extracted from a pilot CEVM PMS to identify patient records with a particular event or term within their record, in this instance we searched for the term CEVM.

(ii) To text mine, validate, and measure the precision of the Keyword in context (KWIC) and Keyword Retrieval (KR) functions in Wordstat to identify patients recorded as having a vaccination consultation from EPRs extracted from a sentinel practice PMS.

3.2 Materials and Methods

For this chapter, IE is described as the extraction of clinical information from veterinary practice management software (PMS) systems and electronic patient records (EPRs). The method of PMS EPR extraction is described in full previously (Chapter 2).

The software system used to support the analysis was a computational content analysis and text mining program WordStat[13] which uses the statistical software SimStat[14] to pre-prepare the data and assist with analysis.

Content analysis describes the pre-processing of an EPR using content analysis software, such as WordStat. Pre-processing may include cleaning or preparing data for further analysis such as selecting all feline patients to create a smaller subset for analysis. Text-mining describes the analysis of an EPR at the text level to identify a word or phrase describing a clinical condition or event. The aim of the analysis was to select a sample of patient records within the smaller subset with a clinical condition common to all. The clinical free-text is analysed to identify words describing the common condition using an

[13] WordStat version 6, Provalis research, Quebec, Canada

[14] SimStat version 2, Provalis research, Quebec, Canada

automated keyword search and retrieval facility within the software text-mining function. The diagram in Figure 3-1 summaries the order of data analysis performed.

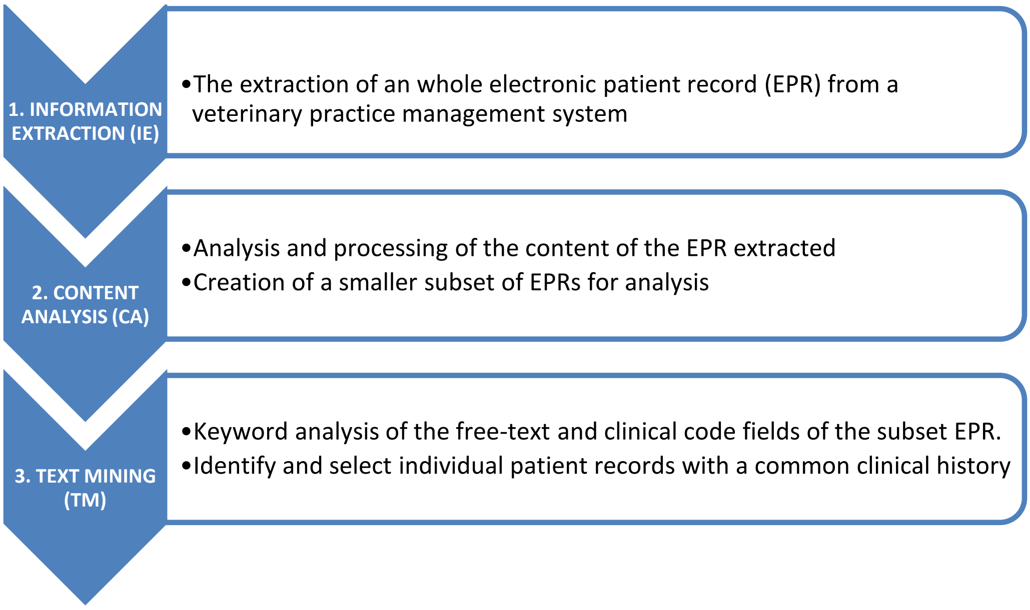


Figure 3-1 A diagram summarising the order of data analysis for electronic patient record extraction.

3.2.1 Electronic Patient Record

The Electronic patient records, extracted from two veterinary PMS systems, were used for data extraction as described in Chapter 2. In brief;

i) A pilot CEVM PMS system. The EPR was populated with (mock) patient data by the researcher JJD.

ii) A working sentinel practice PMS. The EPR was extracted for all small animal consultations during normal working hours over an 8 week period.

The methods of Information extraction (discussed in Chapter 2) and Content analysis were the same for both EPR data sets as follows;

3.2.2 Content Analysis

The patient records were cleaned in Excel by performing a search and removal of unnecessary characters (Appendix 5) and a check of worksheet formatting to ensure the data had been imported into the correct fields. Then the whole dataset was imported into SimStat for analysis. SimStat is a powerful statistics program integral to the WordStat Provalis software package. The Animal ID field, Text Entry field and Diagnosis field were selected for content analysis (Figure 3-2). The Content Analysis function was then selected which opened WordStat into a separate window and imported all data from the selected fields into the WordStat program for text mining.

3.2.3 Text Mining

The Text mining method was designed to identify keywords within the context of the Text Entry and Diagnosis fields of the extracted information. Two methods of analysis were piloted for the Text mining method, 'Keyword in context (KWIC)' and 'Keyword Retrieval (KR)' search facilities.

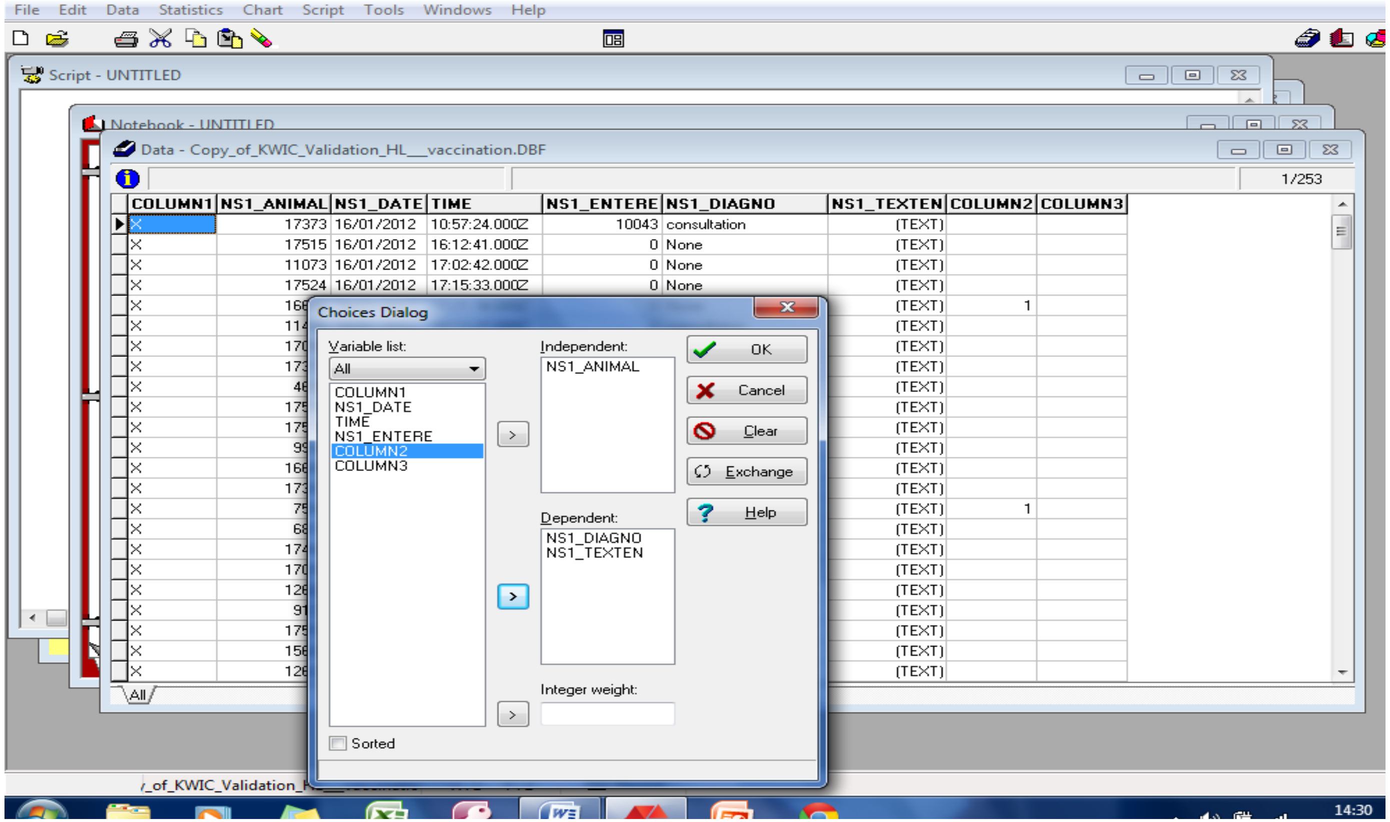


Figure 3-2 Screenshot of content analysis function in SimStat

3.2.3.1 Project 1 Keyword in context (KWIC)

The KWIC search method allowed detection of all instances a defined word occurred within the clinical notes, including the Text Entry and the Diagnosis fields. The results of the search were displayed in a table, in the context of the sentence in which it was located, to allow a judgement on whether the words used met the requirements for inclusion; in this case they identified a vaccination.

3.2.3.1.1 CEVM pilot PMS

The CEVM pilot PMS system was created for the purpose of this research. Therefore a known number of keywords were included in the Text Entry and Diagnosis field during the creation of the EPR. This was to pilot the text mining method and to validate the analysis. The keywords entered into the Text Entry field were '**CEVM**', '**JJD**' '**Vaccination**' and '**Julie Jones-Diette'**. It was decided a single search term, 'CEVM', from those available would be chosen for the search. WordStat is case insensitive, however to confirm this upper and lower case letters were also used during entry of text into the EPRs initially.

3.2.3.1.2 Sentinel PMS

For the sentinel PMS EPR analysis, all patient records from the 10% sample set (n = 252), as described in Chapter 4, were included. The keyword selected for the EPR search was 'vaccination' as this was a term likely to result in a number of positive identifications within the KWIC search and could be easily identified by a manual search. Therefore records where the patient had the term vaccination in their record, or there was charge for a vaccination, in the EPR would be considered a positive result. The term *vac*, with a wildcard search method was used when searching within the EPRs, the asterix either side of the truncated term ensuring that all variations of the word would be identified e.g. vacc, vaccination, vaccines, Nobivac etc. would all be highlighted.

3.2.3.2 Project 2 Keyword Retrieval

3.2.3.2.1 Sentinel PMS

The keyword retrieval (KR) search method allowed an investigation of the EPR using many keywords simultaneously. The sample dataset was as before, all patient records within the 10% sample set (n = 252) as described in Chapter 4, from the sentinel PMS EPR where the patient had received, or a charge was made, for a vaccination. The records were searched using terms selected from a list provided by WordStat frequency analysis function. The frequency analysis creates a list of all lexicon (stand-alone) terms within the language used in the free-text, aside from the grammatical linking terms, and includes an analysis of how many times each term appears. A drag and drop function allows one to select terms from this list to create a smaller specific list with which to search.

The list of keywords selected from the frequency analysis were; vac, vaccination, Svaccination, Nobivac, Lambivac. The term Svaccination was selected from the available list of terms, as the addition of 'S' to the word is a practice specific code used within the clinical notes, in this instance to indicate a specific service provided and which was invoiced.

3.2.4 Validation Analysis

Validation was performed in both the CEVM and Sentinel systems. By comparing the number of patient records correctly identified from the EPR extraction using the WordStat 'KWIC' search, and the 'KR' search, compared to those that had the search term present in their clinical notes. The patient records where a vaccination had been given were identified by a manual search of all records, performed by the researcher and considered the 'gold standard' test for this method.

3.2.5 Statistical Measures

The results from the manual and automated searches were assessed by calculating sensitivity, specificity, positive predictive value (*PPV*) and negative predictive value *(NPV)* using the equations as stated by Petrie and Watson (2006) and described in Appendix 6, and 95% confidence intervals (CI) were calculated, where appropriate, using the methods for proportional values greater than 95% (Wilson, 1927).

3.3 Results

3.3.1 Project 1 - KWIC Search

3.3.1.1 CEVM pilot EPR KWIC Search

Sensitivity, specificity and predictive values were all found to be excellent (Table 3-1). A WordStat KWIC search of the CEVM EPRs, using "CEVM" as the search term retrieved 24 occurrences of the keyword within 23 patient records Table 3-2. One patient record contained the word "cevm" twice, (Figure 3-3). A manual examination of the 326 patient records found there were indeed 23 patient records containing the term "CEVM" or "cevm" (Table 3-2). This resulted in a high sensitivity of 100% and a positive predictive value of 95.8% and a negative predictive value of 100% (Table 3-1).

Table 3-1 Results of the WordStat KWIC search of extracted data from the created CEVM veterinary PMS using "CEVM" as the search term within the [Text Entry] and [Diagnosis] fields.

	Manual count (Gold Standard)		
KWIC Search 'CEVM'	"CEVM" Present	"CEVM" Not present	Total
Positive	23	1	24
Negative	0	302	302
Total	23	303	326

Sensitivity = 23/23 x 100 = 100% (95% CI 85.6% to 100%)

Specificity = 302/303 x 100 = 99.7% (95% CI 98.1% to 99.9%)

PPV = 23/24 x 100 = 95.8% (95% CI 97.9% to 99.2%)

NPV = 302/302 = 100% (95% CI 98.7% to 99.9%)

Table 3-2 List of records from the CEVM system with 'CEVM' in the clinical notes.

NUMBER OF OCCURANCES	NS1_ANIMAL ID	VARIABLE	KEYWORD	NS1_SPECIES	NS1_BREED
1	1784534	NS1_TEXTEN	CEVM	Cat	Domestic Short Hair
2	1784535	NS1_TEXTEN	CEVM	Cat	Domestic Short Hair
3	1784536	NS1_TEXTEN	CEVM	Cat	Birman
4	1784538	NS1_TEXTEN	CEVM	Cat	DSH
5	1784543	NS1_TEXTEN	CEVM	Cat	
6	1784547	NS1_TEXTEN	CEVM	Cat	Domestic Short Hair
7	1784551*	NS1_TEXTEN	CEVM	Cat	Persian
8	1784551*	NS1_TEXTEN	cevm	Cat	Persian
9	1784556	NS1_TEXTEN	CEVM	Cat	Domestic Short Hair
10	1784564	NS1_TEXTEN	CEVM	Dog	Cross Breed - Small
11	1784565	NS1_TEXTEN	cevm	Dog	Cross Breed - Medium
12	1784568	NS1_TEXTEN	CEVM	Dog	Cross Breed - Small
13	1784571	NS1_TEXTEN	CEVM	Dog	Cross Breed - Small
14	1784576	NS1_TEXTEN	cevm	Dog	Cross Breed - Medium
15	1784582	NS1_TEXTEN	CEVM	Dog	Cross Breed - Small
16	1784583	NS1_TEXTEN	cevm	Dog	Bearded Collie
17	1784584	NS1_TEXTEN	CEVM	Dog	Boxer
18	1784589	NS1_TEXTEN	CEVM	Dog	Cross Breed - Small
19	1784601	NS1_TEXTEN	CEVM	Dog	Cross Breed - Small
20	1784603	NS1_TEXTEN	cevm	Dog	Bearded Collie
21	1784605	NS1_TEXTEN	CEVM	Dog	Cross Breed - Small
22	1784608	NS1_TEXTEN	CEVM	Cat	DLH
23	1784609	NS1_TEXTEN	CEVM	Dog	Boxer
24	1784612	NS1_TEXTEN	cevm	Dog	Bedlington Terrier
n = 24	***One animal had the keyword twice within their patient record**				

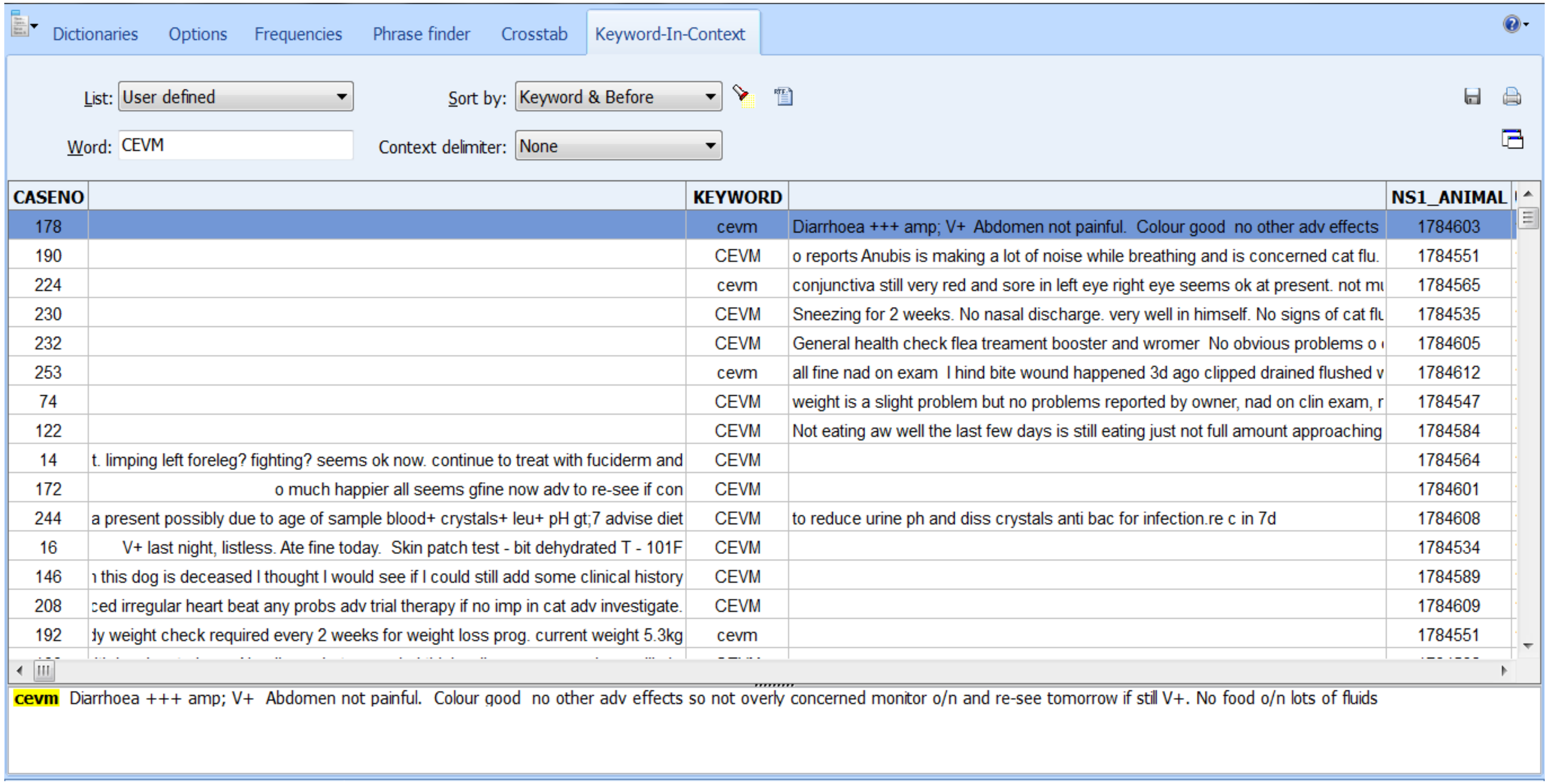


Figure 3-3 Keyword in context search of CEVM PMS system.

3.3.1.2 Sentinel EPR KWIC Search

The final total number of patient records identified as having a vaccination was 29 animals (*PPV 95.8* %), Table 3-3. The KWIC search of the sentinel EPR using the search term *vac* (Figure 3-4) found 31 occurrences of variations of vac within patient records. However on two occasions a single patient record had two occurrences of vac within the same patient record (Table 3-4).

In the sentinel EPR, the manual count (gold standard measure for this method) identified 34 patient records where a vaccination was given to the animal or charged to the client. The recall of the KWIC search method therefore using the sentinel EPR was good but missed 5 positive records producing a Sensitivity of 85.3% (Table 3-3). However the precision was still very good, with a PPV = 93.5% and NPV = 97.7%.

Table 3-3 Results (observed frequencies) of the WordStat KWIC search of extracted data from the sentinel veterinary PMS using *vac* as the search term

	Manual count (Gold Standard)		
KWIC Search *vac*	*vac* Positive	*vac* Negative	Total
Positive	29	2	31
Negative	5	216	221
Total	34	218	252

Sensitivity = 29/34 x 100= 85.3% (95% CI 69.8% to 93.5%)

Specificity = 216/218 x 100 = 99.1% (95% CI 96.7% to 99.7%)

PPV 29/31 x 100 = 93.5% (95% CI 82.0% to 100%)

NPP 216/221 x 100 = 97.7% (95% CI 94.8% to 99.0%)

Observed prevalence = 34/252 = 14%

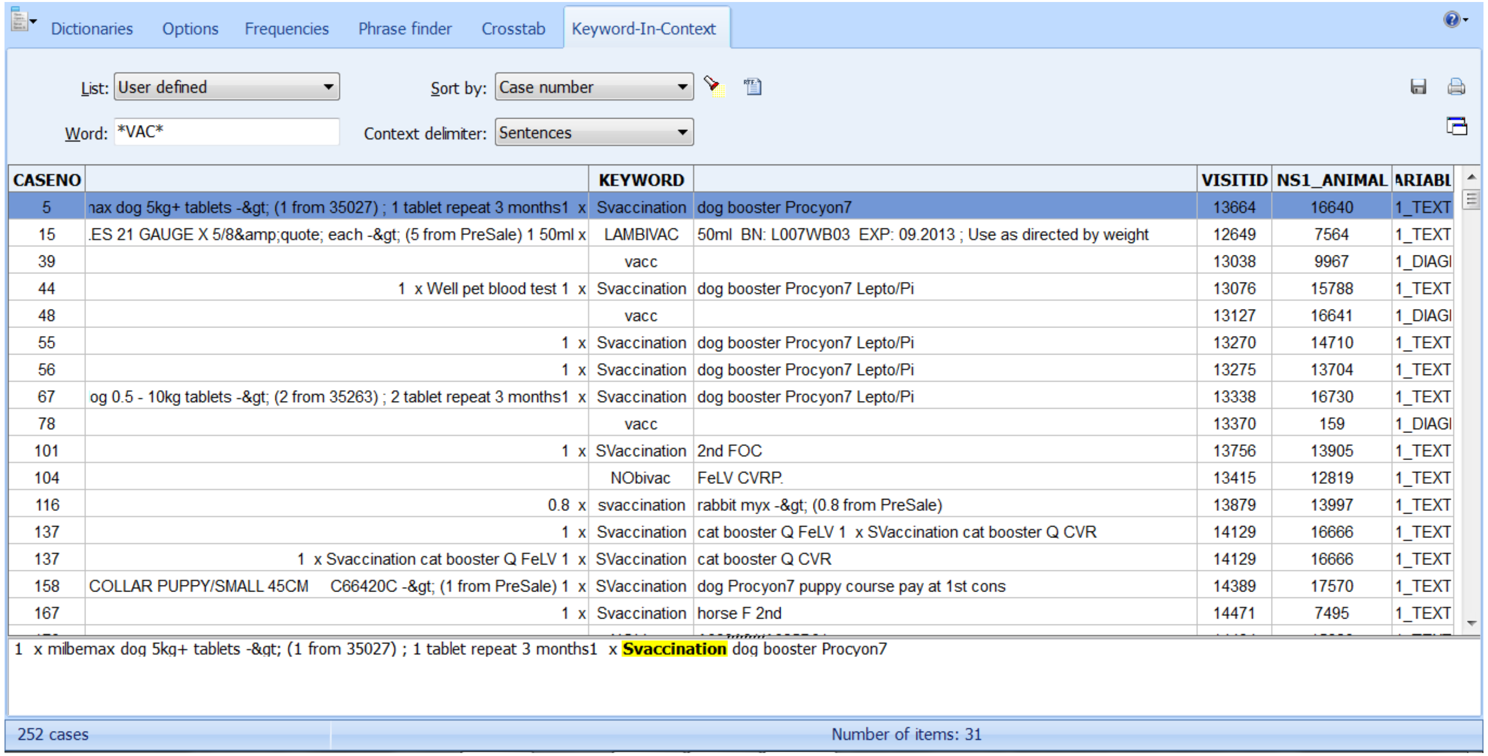


Figure 3-4 Example of the output screen for the KWIC search using *vac* as the wildcard search term.

Table 3-4 Sentences of text identified by WordStat KWIC search (n=31) and duplicate records highlighted where the term search term occurred twice (Animal ID 16666 and Animal ID 9217).

VISITID	NS1_ANIMAL	VARIABLE	CLINICAL NOTES	KEYWORD	CLINICAL NOTES
13664	16640	NS1_TEXTEN	1 x milbemax dog 5kg+ tablets - > (1 from 35027) ; 1 tablet repeat 3 months1 x	Svaccination	dog booster Procyon7
12649	7564	NS1_TEXTEN	5 Each x NEEDLES 21 GAUGE X 5/8&quote; each -> (5 from PreSale) 1 50ml x	LAMBIVAC	50ml BN: L007WB03 EXP: 09.2013 ; Use as directed by weight
13038	9967	NS1_DIAGNO		vacc	
13076	15788	NS1_TEXTEN	1 x Well pet blood test 1 x	Svaccination	dog booster Procyon7 Lepto/Pi
13127	16641	NS1_DIAGNO		vacc	
13270	14710	NS1_TEXTEN	1 x	Svaccination	dog booster Procyon7 Lepto/Pi
13275	13704	NS1_TEXTEN	1 x	Svaccination	dog booster Procyon7 Lepto/Pi
13338	16730	NS1_TEXTEN	1 chip x TRACER CHIP -> (1 from PreSale) 2 x Milbemax dog 0.5 - 10kg tablets -> (2 from 35263) ; 2 tablet repeat 3 months1 x	Svaccination	dog booster Procyon7 Lepto/Pi
13370	159	NS1_DIAGNO		vacc	
13756	13905	NS1_TEXTEN	1 x	SVaccination	2nd FOC
13415	12819	NS1_TEXTEN		NObivac	FeLV CVRP.
13879	13997	NS1_TEXTEN	0.8 x	svaccination	rabbit myx -> (0.8 from PreSale)
14129	16666	NS1_TEXTEN	1 x	Svaccination	cat booster Q FeLV 1 x SVaccination cat booster Q CVR
14129	16666	NS1_TEXTEN	1 x Svaccination cat booster Q FeLV 1 x	SVaccination	cat booster Q CVR
14389	17570	NS1_TEXTEN	1 x Advocate 40mg < 4KG free puppy -> (1 from PreSale) ; On skin repeat monthly1 x ADAPTIL COLLAR PUPPY/SMALL 45CM C66420C -> (1 from PreSale) 1 x	SVaccination	dog Procyon7 puppy course pay at 1st cons
14471	7495	NS1_TEXTEN	1 x	Svaccination	horse F 2nd
14484	15820	NS1_TEXTEN		NObivac	A089A01/A025B01.
14549	16664	NS1_TEXTEN		Nobivac	PIL A025A01/A089A01.
14539	17579	NS1_TEXTEN	1 x Advocate 100mg 4 - 10 KG free puppy -> (1 from PreSale) 1 x	SVaccination	dog Procyon7 2nd
14776	12768	NS1_TEXTEN	1 x	Svaccination	dog booster Procyon7 1 x ADVOCATE SP/ON DOG LARGE 10-25Kg 6 PIP 250 - > (1 from KP07FRD) ; On skin repeat monthly1 x MILBEMAX REMINDER 3 MONTHS -> (1 from PreSale) 3 x milbemax dog 5kg+ tablets -> (3 from 35121) ; Give 1 tablet as a si
14657	12922	NS1_TEXTEN	1 x	Svaccination	dog booster Procyon7 Lepto/Pi
14869	9217	NS1_TEXTEN	1 x	SVaccination	Kennel cough 1 x Svaccination dog booster Procyon7 Lepto/Pi
14869	9217	NS1_TEXTEN	1 x SVaccination Kennel cough 1 x	Svaccination	dog booster Procyon7 Lepto/Pi
14951	11932	NS1_TEXTEN	1 x	Svaccination	horse F 3rd 1 x Svisit 1
14943	12481	NS1_TEXTEN	1 x	SVaccination	dog Procyon7 1st Amnesty
15041	17598	NS1_DIAGNO		vacc	
15043	17244	NS1_TEXTEN	1 x	SVaccination	2nd FOC
15044	17245	NS1_TEXTEN	1 x	SVaccination	2nd FOC
15069	16228	NS1_TEXTEN		Nobivac	PIL A026A01/ A087A01.
15075	3030	NS1_TEXTEN	4 x milbemax dog 5kg+ tablets - > (4 from 35639) ; 2 tablets per dog repeat 3 months1 x	Svaccination	dog booster Procyon7
15083	14911	NS1_TEXTEN	1 x MILBEMAX REMINDER 3 MONTHS -> (1 from PreSale) 1 x Milbemax dog chewy tabs 5kg+ -> (1 from PreSale) ; 1 tablet repeat 3 months1 x Well pet blood test 1 x	Svaccination	dog booster Procyon7 Lepto/Pi 1 x Svisit with nurse

Field Key: Visit ID as allocated by the Access data warehouse, Animal ID allocated by the PMS system, Variable describes the field where the term was found NS1_ TEXT refers to the clinical notes, NS1_DIAGNO refers to the clinical code recorded within the diagnosis field.

3.3.2 Project 2 - Keyword Retrieval

3.3.2.1 Sentinel EPR Keyword Retrieval Search

The keyword retrieval search focussed on records containing any one of the list of keywords selected. The search identified 29 out of a possible 34 patient records where a vaccination had been recorded in the animals EPR, so missed 5 records. The search also identified 218 out of a possible 218 where the animals had no record of a vaccination in their EPR (Table 3-5). The results of the Keyword Retrieval search of the sentinel EPR therefore was found to have a slightly higher PPV (100%) than the KWIC search (PPV=94%) using the same sentinel dataset to identify patient records with one of the keywords present.

The keyword retrieval search method missed 5 patient records (false negatives) which should have been identified as positive records as confirmed by the manual search where a reference to vaccination was present in their record. The records were found to have no mention of the term ‘vaccination’ or even ‘vac’ in their EPR and were only identified due to the input of a vaccine batch number by the veterinarian into the animals EPR which was recognised during the manual count.

Table 3-5 Results (observed frequencies) of the WordStat Keyword Retrieval search of extracted data from the sentinel veterinary PMS using BOOSTER, LAMBIVAC, NOBIVAC, SVACCINATION and VACC as the search terms

	Manual count "Gold Standard"		
Keyword Retrieval Search of patient record	Vaccination present in clinical record	Vaccination not present in clinical record	Total
Term present	29	0	29
Term not present	5	218	223
Total retrieved	34	218	252

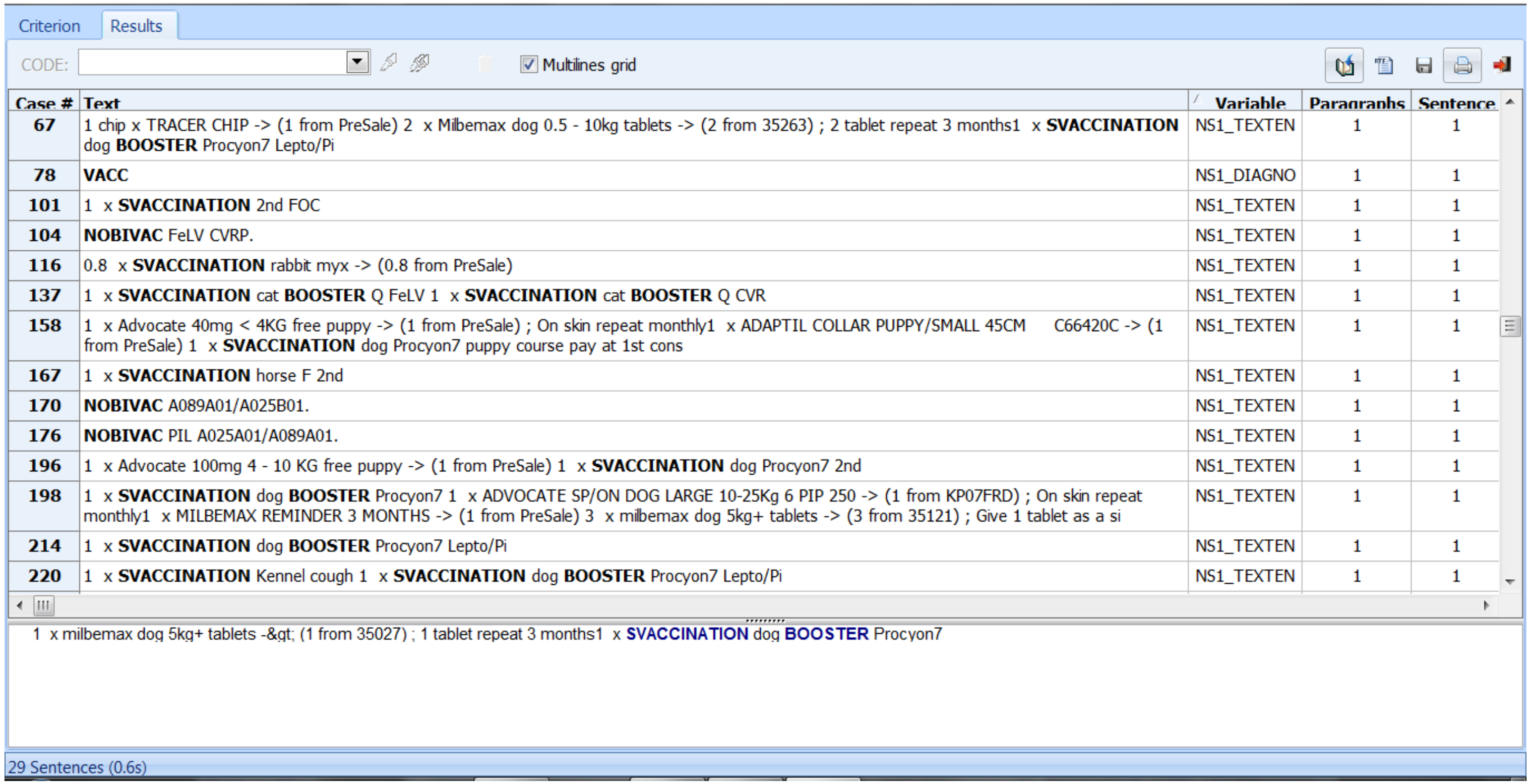
Sensitivity = 29/34 x 100 = 85.3% (95% CI 69.8% to 93.5%)

Specificity = 218/218 x 100 = 100% (95% CI 98.2% to 100%)

PPV 29/29 x 100 = 100% (95% CI 88.0% to 100%)

NPP 218/223 x 100 = 97.7% (95% CI 94.8% to 99.0%)

Observed prevalence = 34/252 = 14%



Case #	Text	Variable	Paragraphs	Sentence
67	1 chip x TRACER CHIP -> (1 from PreSale) 2 x Milbemax dog 0.5 - 10kg tablets -> (2 from 35263) ; 2 tablet repeat 3 months1 x **SVACCINATION** dog **BOOSTER** Procyon7 Lepto/Pi	NS1_TEXTEN	1	1
78	**VACC**	NS1_DIAGNO	1	1
101	1 x **SVACCINATION** 2nd FOC	NS1_TEXTEN	1	1
104	**NOBIVAC** FeLV CVRP.	NS1_TEXTEN	1	1
116	0.8 x **SVACCINATION** rabbit myx -> (0.8 from PreSale)	NS1_TEXTEN	1	1
137	1 x **SVACCINATION** cat **BOOSTER** Q FeLV 1 x **SVACCINATION** cat **BOOSTER** Q CVR	NS1_TEXTEN	1	1
158	1 x Advocate 40mg < 4KG free puppy -> (1 from PreSale) ; On skin repeat monthly1 x ADAPTIL COLLAR PUPPY/SMALL 45CM C66420C -> (1 from PreSale) 1 x **SVACCINATION** dog Procyon7 puppy course pay at 1st cons	NS1_TEXTEN	1	1
167	1 x **SVACCINATION** horse F 2nd	NS1_TEXTEN	1	1
170	**NOBIVAC** A089A01/A025B01.	NS1_TEXTEN	1	1
176	**NOBIVAC** PIL A025A01/A089A01.	NS1_TEXTEN	1	1
196	1 x Advocate 100mg 4 - 10 KG free puppy -> (1 from PreSale) 1 x **SVACCINATION** dog Procyon7 2nd	NS1_TEXTEN	1	1
198	1 x **SVACCINATION** dog **BOOSTER** Procyon7 1 x ADVOCATE SP/ON DOG LARGE 10-25Kg 6 PIP 250 -> (1 from KP07FRD) ; On skin repeat monthly1 x MILBEMAX REMINDER 3 MONTHS -> (1 from PreSale) 3 x milbemax dog 5kg+ tablets -> (3 from 35121) ; Give 1 tablet as a si	NS1_TEXTEN	1	1
214	1 x **SVACCINATION** dog **BOOSTER** Procyon7 Lepto/Pi	NS1_TEXTEN	1	1
220	1 x **SVACCINATION** Kennel cough 1 x **SVACCINATION** dog **BOOSTER** Procyon7 Lepto/Pi	NS1_TEXTEN	1	1

Figure 3-5 Output screen for results of Keyword retrieval search criteria used for search of sentinel practice PMS system.

3.4 Discussion

The Wordstat program was found to be an excellent resource for the content analysis of extracted EPRs using text-mining techniques. The Keyword in context (KWIC) search was sufficient to identify almost all positive cases and the precision was high enough to provide a useful assessment of the dataset. However, keyword retrieval was found to be superior in this circumstance as it avoided false positive results allowing for greater precision overall.

3.4.1 Project 1 – Keyword in context

The KWIC search of the CEVM pilot EPR was a useful pilot test of the software. The method worked well. All the patients who were identified did indeed have the 'CEVM' term within their notes. The duplicate count of one patient in this instance reduced the sensitivity of the method which could be a concern over a larger scale dataset, leading to an over estimation of prevalence within a real situation. However it is possibly better to overestimate than underestimate disease prevalence.

The sentinel EPR offered the opportunity to test the KWIC method on real patient data which contained natural variation in terminology. The results of this analysis showed a good level of accuracy between the actual numbers of positive patient records within the EPR and those found using the method. The method also suggested a high level of sensitivity for the test. However, this was affected somewhat by the number of cases missed owing to the recording of vaccinations by batch code alone on the part of the veterinary team. The use of only a batch code to record a vaccination in the EPR is a new finding and highlights the need for greater care in recording of information but also allows for development of the method. Although the batch code was used to record the administration of the vaccine, the invoice will record a vaccination was given therefore by checking both the free text and the invoice records both forms of data would be captured.

Prevalence and incidence estimates in clinical epidemiology have a degree of trade-off between sensitivity and specificity and it is less of a concern if false positives are identified as the consequence may simply be a slight over estimation. This is in contrast to diagnostic tests where a true positive is essential if the disease is present for further action to be taken or disastrous if a false positive is recorded e.g. in the case of Foot and Mouth. Ultimately it depends on the purpose of the test and how the results will be used.

Sensitivity and specificity measure how accurate the methods are and are a measure of reliability of the test. Predictive values measure how useful the methods are at identifying cases within the assessed population (Petrie and Watson, 2006) and as such are dependent on prevalence of the disease within that population. Positive predictive values will be higher where the disease is common within a population and conversely negative predictive values will be lower. For the purpose of practice-based research such as this it is important true positives for disease are identified (Petrie and Watson, 2006), particularly if the purpose is to assist with treatment decisions. Therefore a high level of sensitivity and positive predictive values were selected as the primary measure for this analysis to ensure the aims of the method were met and as many patients as possible with the keyword present in their record (true positives) could be identified accurately and reliably. Other measures were included to allow for a greater understanding of the method success as a whole.

Singular KWIC searching has its limitations especially if the single word used for the search is not specific enough to identify the intended records. Nevertheless, the specificity of the KWIC method was very high for analysis of the sentinel EPR. This can be explained by the fact the method rarely identifies a true negative record as a false negative; the results suggest provided the keyword of interest is present in the patient record the KWIC search should be able to locate it.

3.4.2 Project 2 - Keyword Retrieval

3.4.2.1 Sentinel PMS Keyword retrieval search of the EPR

The Keyword Retrieval search allows for a group of terms to use in the search and therefore has the potential for greater capture of data. Keyword retrieval method had excellent sensitivity, due in part to the option to select search terms from a list created by the WordStat program using the vaccination terms or codes used by the veterinarians themselves within the EPR, assuming they are recognisable as a vaccination term.

There were five patient records from the sentinel practice PMS which were not identified by either the KWIC search or Keyword Retrieval. It is likely this was due to the vaccination recording method (e.g. batch number only) used by the sentinel practice. Understandably these could not be identified by either search method without prior knowledge of batch code information unless the invoice had been examined in addition to the free text of the consultation, and for this study was not the case. All records were selected using the random sampling method and the randomisation selected a single patient record rather than all details of a single visit in the EPR. Each patient selected at random had either a consultation record or an invoice record by the randomisation but not both. As such it was very likely the 5 records missed during the search would have been identified if the full dataset (n = 2519) had been used or the patients data had been selected as a whole. This is a limitation of the study design but not the methodology. A recommendation for the study design therefore is that where a sample of records are chosen, both the free text field in the consultation record and the invoice record for the consultation should be selected together for analysis.

Sensitivity analysis used in this research allowed for a measure of diagnostic/search ability but it is the predictive values that indicate how likely it is that the results are representative of the larger sample or population

Petrie and Sabin (2009). The *PPV* for all of the analyses were found to be very high, ranging from 93-100%. This is an excellent result and allows confidence in the use of the WordStat search function for identifying vaccination records at least, particularly the Keyword Retrieval method which produced a *PPV* of 100%. The relatively high *PPV* of the keyword retrieval method may be a function of the observed prevalence of vaccination consultations within the sentinel EPR which was quite a high proportion at 14% (Estberg et al., 1998).

As already mentioned the sensitivity analysis was used preferentially as a measure of accuracy for our test to identify disease prevalence in the sample population and particularly in the case of the data collected from the sentinel practice where validation of the actual prevalence could not be quantified. For the next steps of this work greater reliance will be placed on the predictive ability and reliability of the test as future research will be a test of performance using many different datasets with more prior understanding of disease prevalence and therefore where measures of PPV and NPV will become much more valid and necessary.

One limitation of the study was the difficulty of keyword retrieval when searching the narrative used by veterinarians in first opinion veterinary practice. As a standard veterinary terminology or clinical coding system is not currently available for use within the profession (Case et al., 2000) it is difficult to anticipate the many different terms vets may use to record clinical signs or even diagnoses.

The veterinarians in the sentinel practice for this study have their own list of clinical codes which they refer to and use within their notes. However the majority of patient records (80%) were coded as “consultation” which meant most of the clinical data was recorded within the Text Entry field. In addition, for invoice purposes the keyword retrieval results highlighted that the sentinel practice used a separate coding system to charge for services placing an S ahead of the item to be charged, eg. Svaccination, Sconsultation and so on. It is also common for busy vets in practice to use a type of short hand for certain terms such as dt for Diet, op for Operation, dx for diagnosis (see Appendix 7).

The WordStat keyword retrieval method was able to overcome some of these complications by allowing the user to select search terms from the text extracted which produced a very accurate result. Although all possible cases could not be identified, it was a very small percentage which was missed (2%) and in fact better than similar work published elsewhere (Lam et al., 2007a). This study suggests the automated method of disease reporting and patient record aggregation is a powerful tool for evidence-based veterinary medicine and research (Lund et al., 1999, Case et al., 2000, Moore et al., 2005, Moore et al., 2007, Johnson et al., 2011, Santamaria and Zimmerman, 2011). The software can provide large scale analysis in a very short time period and far exceeds the ability of human search and retrieve capability within the same timeframe. The results of the study suggest the program would be an excellent resource for practice-based research using many PMS systems and extracted EPRs.

3.5 Conclusion

The methods presented here suggest the content analysis and text mining software used, particularly the Keyword Retrieval function, provided a high level of precision for search and recall of patient records sharing common clinical information. However it must be noted with all of these results that conclusions can only be drawn on the search and find accuracy for records where the veterinarians have recorded a vaccination, populations with more complex cases of disease with a more unpredictable terminology may provide more of a challenge.

Chapter Four

4 Mining clinical data extracted from veterinary practice

4.1 Introduction

To be able to estimate the prevalence and incidence of disease, examine treatment outcomes and the result of clinical decisions within the vet visiting pet population, it is necessary to not only gain access to the animal's clinical record but also be able to analyse the data retrieved. To identify patient records where a similar clinical condition has been recorded, text mining has been found to be a useful method in both medical (Hersh, 2003b, Meystre et al., 2008, Tate et al., 2011, Nicholson et al., 2011) and veterinary research (Estberg et al., 1998, Lam et al., 2007a, Lam et al., 2007b, Lam et al., 2007c). However one must first understand the language a veterinarian may use to record clinical signs or diagnoses before one can search for their presence in the EPR. Therefore a challenge for practice-based research, when using veterinary electronic patient records to gather information, is interpretation of the clinical narrative (Szolovits, 2003, Batal and Hauskrecht, 2010). The lack of a standardised medical terminology within the veterinary profession adds to the complexity of this issue (Lund, 1997, Lund et al., 1998, Case et al., 2000, Faunt et al., 2007).

Information retrieved from veterinary practice medical records suggests the terminology used by veterinarians to record animal clinical signs may differ widely between different practices (Lund et al., 1998) and between veterinarians within the same practice (Estberg et al., 1998). Synonyms, abbreviations, shorthand, semantic nuances and colloquialism may all complicate the process of disease classification during text mining (Hersh, 2003b, Meystre et al., 2008). In addition veterinarians may use unique terms learnt from mentors, peers or taught during teaching and training (Appendix 7). The extensiveness of the medical vocabulary used can also influence the success of extracting information from veterinary clinical data (Lund et al., 1998, Estberg et al., 1998, Case et al., 2000, Faunt et al., 2007).

The purpose of the research presented here was to develop a text mining methodology, including a dictionary of disease specific veterinary language, to analyse electronic patient records extracted from a first opinion sentinel veterinary practice. A common clinical condition; feline lower urinary tract disease (FLUTD), was chosen to trial the methodology (Allen and Kruger, 2000).

4.2 Materials and Methods

4.2.1 Study Design

Electronic patient records (EPRs) extracted from the veterinary practice management software system (PMS) of a single sentinel practice were used to develop a text mining methodology. The efficacy of a computer assisted free text analysis program (WordStat) was assessed, for the analysis of textual information recorded in the EPR, using a dictionary of clinical terms. Method precision was determined using sensitivity, specificity and positive predictive value analysis using the methods described by Petrie and Watson (2006), see Appendix 6 for equation details.

4.2.2 Sample Population

The population data used for the analysis was patient records extracted from a sentinel first opinion veterinary practice over an 8 week period, 16th Jan – 11th March 2012 (see Chapter 2 for details). The clinical condition chosen to test the text mining search performance was feline lower urinary tract disease (FLUTD).

To identify and select out the population for further analysis all records were filtered by animal species, selecting feline only, and all associated data was retrieved and a new file created (Figure 4-1). Once all the feline records were retrieved, the invoice records were excluded from the analysis as they are often a duplication of the consultation information. Identification of the invoice information was possible as the sentinel vet practice coded their invoice records within the diagnosis field of the EPR as ‘None’.

The fields selected from the EPR included;

- Animal ID
- Species
- Breed
- Date of Birth
- Gender
- Date and time of consultation
- Veterinary ID
- Text entry field
- Diagnostic code

4.2.3 Veterinary Review – The Gold Standard

To compare the effectiveness of the text mining analysis, all feline records were manually searched by a team of qualified veterinarians given full instructions on what was required and how to report their findings to ensure their reports were consistent and comparable. This was the gold standard measure for this study.

The 286 individual feline patient records were numbered and then randomised using a Microsoft Excel randomisation function. To quantify the number of patient records containing words that are consistent with FLUTD terminology a team of five veterinarians were asked to review the data. They were asked to identify records where there were any terms they thought related to a diagnosis of or clinical signs suggestive of FLUTD. The veterinarian reviewers were allocated records each to read and review. The positive cases were identified and highlighted by the veterinarian reviewer within the dataset. The researcher (JJD) remained blinded to the number of records highlighted until after the text mining analysis had been performed.

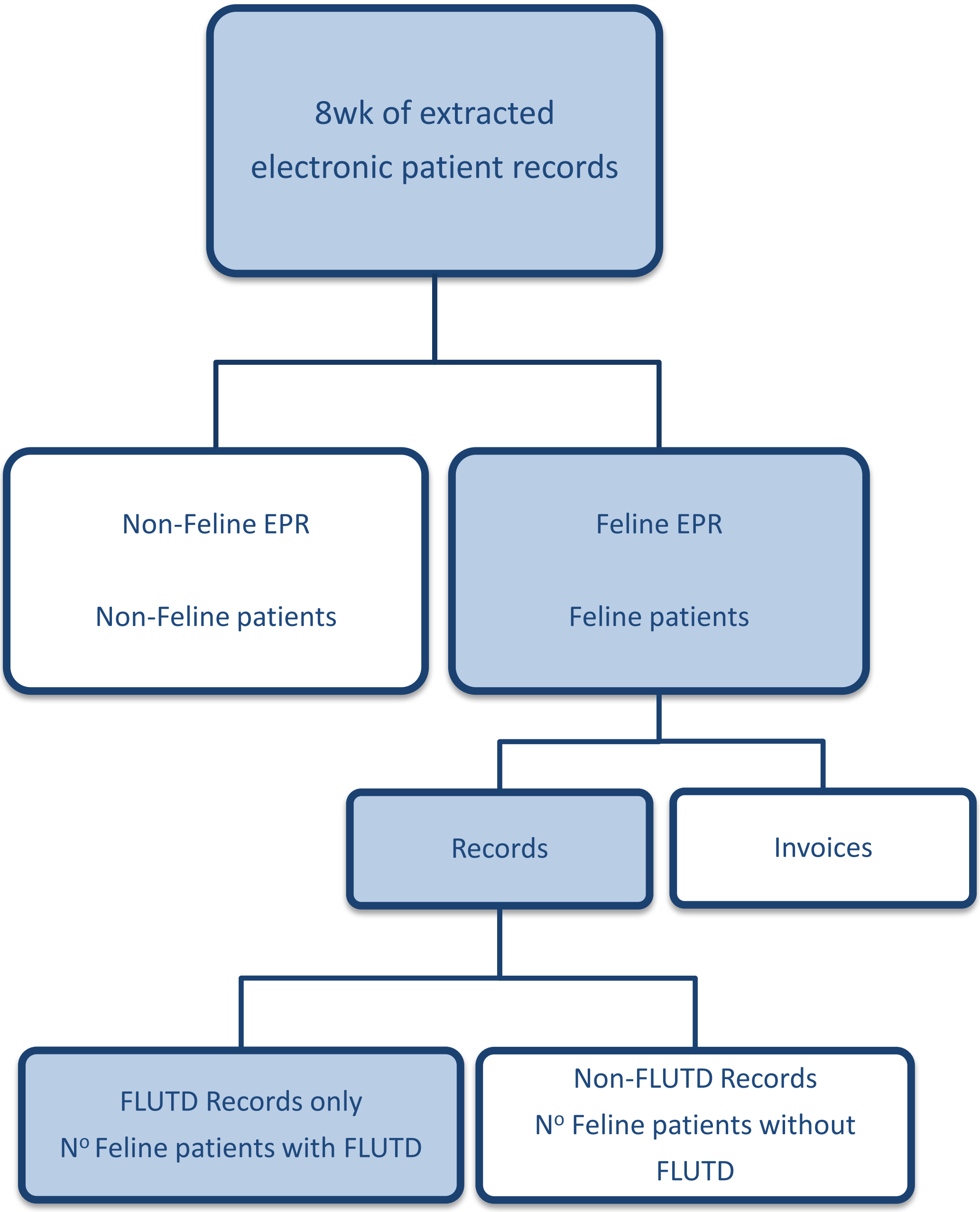


Figure 4-1 Nested target diagram to show selection process for study population

4.2.4 Dictionary 1.0 – Survey Data

The initial text mining dictionary was created by the researcher and was composed of an alphabetical list of diagnostic terms and clinical signs related to FLUTD. The initial dictionary was compiled from data collected as part of a survey of the veterinary profession (Nielsen et al., 2014) in the UK undertaken in 2010. The survey was distributed to all members of the Royal College of Veterinary Surgeons (RCVS) (n = 14532) who agreed to be contacted by third parties for the purposes of research. The response rate for the survey was 33.3% (n=4842).

Veterinarians were asked to detail conditions or complaints they most often see in clinical practice. The responses were collated and categorised according to affected body system or topic. To collate information for the dictionary, all survey responses with reference to the term 'cat' and 'flutd' were selected, including any conditions of the bladder, and was used to form the basis of a clinical dictionary of terms which may describe FLUTD (Table 4-1). The survey and terms selected (n = 123 terms or phrases) were transferred into the dictionary exactly as they were written into the survey including any misspellings, semantics, synonyms and nuances. The initial dictionary was termed FLUTD Dictionary 1.0.

Table 4-1 Terms used by veterinarians in private practice to describe conditions commonly seen in cats for flutd and/or conditions of the bladder (n = 4842 veterinarian respondents). These terms formed the basis of Text Mining Dictionary 1.0.

Diagnosis or Condition	Count	Diagnosis or Condition	Count
cystitis	234	urine infection	2
flutd	211	urinary obstruction	2
uti	78	Bladder infections	2
Urinary Tract Disease	73	Feline lower urinary tract disease	2
urinary problems	58	urinary problem	2
urinary	50	urinary tract disorder	2
urinary tract	39	UTI's	2
urinary tract infections	36	urine problems	2
urinary tract problems	27	blocked cats	2
urinary tract infection	26	Urologic	2
Lower urinary tract disease	20	urinary tract problem	2
Urinary disease	16	urinary tract dx	2
urinary infections	15	Urinary Disorders	2
fus	13	urinary dz	2
urinary tract inf	9		
urinary tract dz	9	Urological	1
UTIs	8	Lower urinary tract dx	1
urinary infection	6	lower urinary tract disaese	1
urinary tract probs	6	feline lower urinary tract	1
urinary tract dis	5	LUT probs	1
urinary conditions	5	urinary blockage	1
blocked bladder	5	uti (blocked cats)	1
urinary tract infec	5	urinary condition	1
urinary probs	5	urinary tract ingection	1
bladder	4	Crystalluria	1
urolithiasis	4	LUT disease	1
dysuria	4	uro genital	1
urinary tract conditions	4	feline lower urinary tract disorder	1
urinary issues	3	urinary disorder	1
Feline idiopathic cystitis	3	ut dx	1
LUTI	3	feline lower urinary tract dx	1
Urinary tract disorders	3	urethral obstructions	1
idiopathic cystitis	3	feline lower urinary tract dz	1
urinary tract obstruction	3	urinary tract inflammatory	1

Diagnosis or Condition	Count	Diagnosis or Condition	Count
ureteral obstruction	1	feline obstructive lower urinary tract dx	1
crystaluria	1	urinary tract intections	1
Urinary tract infect	1	Feline Urinary problems	1
Urinary signs	1	lower urinary tract dz	1
blocked urethra	1	urinary infns	1
urinary straining	1	urinatry tract disease	1
urinary tract infn	1	feline urinary syndrome	1
urinary symptoms	1	LUTD	1
urinary tract ingections	1	feline urinary tract disease	1
urinary tenesmus	1	blocked urinary tract	1
urinary tract issues	1	urinary prob	1
Flutd urinary tract	1	Urology	1
urinary tract prob	1	FIC	1
Urinary tract abnormalities	1	Utd	1
lu tract infections	1	FLUT	1
urinary tract blockage	1	FLUTI	1
urinary treact issues	1	urine infections	1
flutd/FIC	1	bladder problems	1
urine	1	urine tract disease	1
lower urinary tract conditions	1	urinary tract diseases	1
urinary tract infecs	1	Urogenital	1
Urinary tract disese	1	UT problems	1
recurrent cystitis	1	Lower Urinary disease	1
Genitourinary	1	stress cystitis	1
urological syndrome	1	urinary tract in	1
cystits	1	UTI.	1
ut disease	1	lower urinary tract	1
intestinal cystitis	1	urethral obstruction	1
		Individual terms/phrases	**123**
		Grand Total	**1096**

4.2.5 Dictionary 1.1 - Text Mined Data

The feline sample population data was imported into SimStat version 2 for windows[14] a statistical package which was used in conjunction with WordStat version 6 for windows[13] to organise the data prior to analysis (ref Chapter 3).

By designing specific dictionaries based on veterinary terminology for clinical signs of disease the program can streamline the process of content analysis of large sets of unstructured veterinary information, clinical notes or veterinary narrative. Once the patient information was imported into WordStat a new text mining dictionary could be created or an existing dictionary developed further by using the drag and drop facility within the WordStat program using terms used by the veterinarian and extracted from the EPR.

The dictionary designed by the researcher was integrated into Wordstat (dictionary 1.0) then WordStat was used to perform a frequency analysis of words found in the feline EPR. The frequency analysis for this study produced a list of ‘leftover words’, from the EPR (n = 1875 words). Wordstat also presented the total word count (n = 20211 words).

The list of leftover words (n=1875 words) was then examined to identify any additional terms which may relate to FLUTD Figure 4-2. The additional words selected from the leftover words within the feline EPR were; Feliway, pH, pupd, stone, straining, ultrasound, urinalysis, urinate, urinating, In combination with Dictionary 1.0 the above additional 9 terms were used to create Dictionary 1.1 which now contained 132 terms or phrases in total.

4.2.6 Statistical Measures

To evaluate the precision of the WordStat dictionary function to identify positive FLUTD patient records, the sensitivity, specificity and positive and negative predictive value (*PPV and NPV*) of search performance was calculated using the equations as stated by Petrie and Watson (2006) and described in Appendix 6. Ninety five percent confidence intervals (95% CI) were calculated, where appropriate, using the methods described for proportional values greater than 95% (Wilson, 1927).

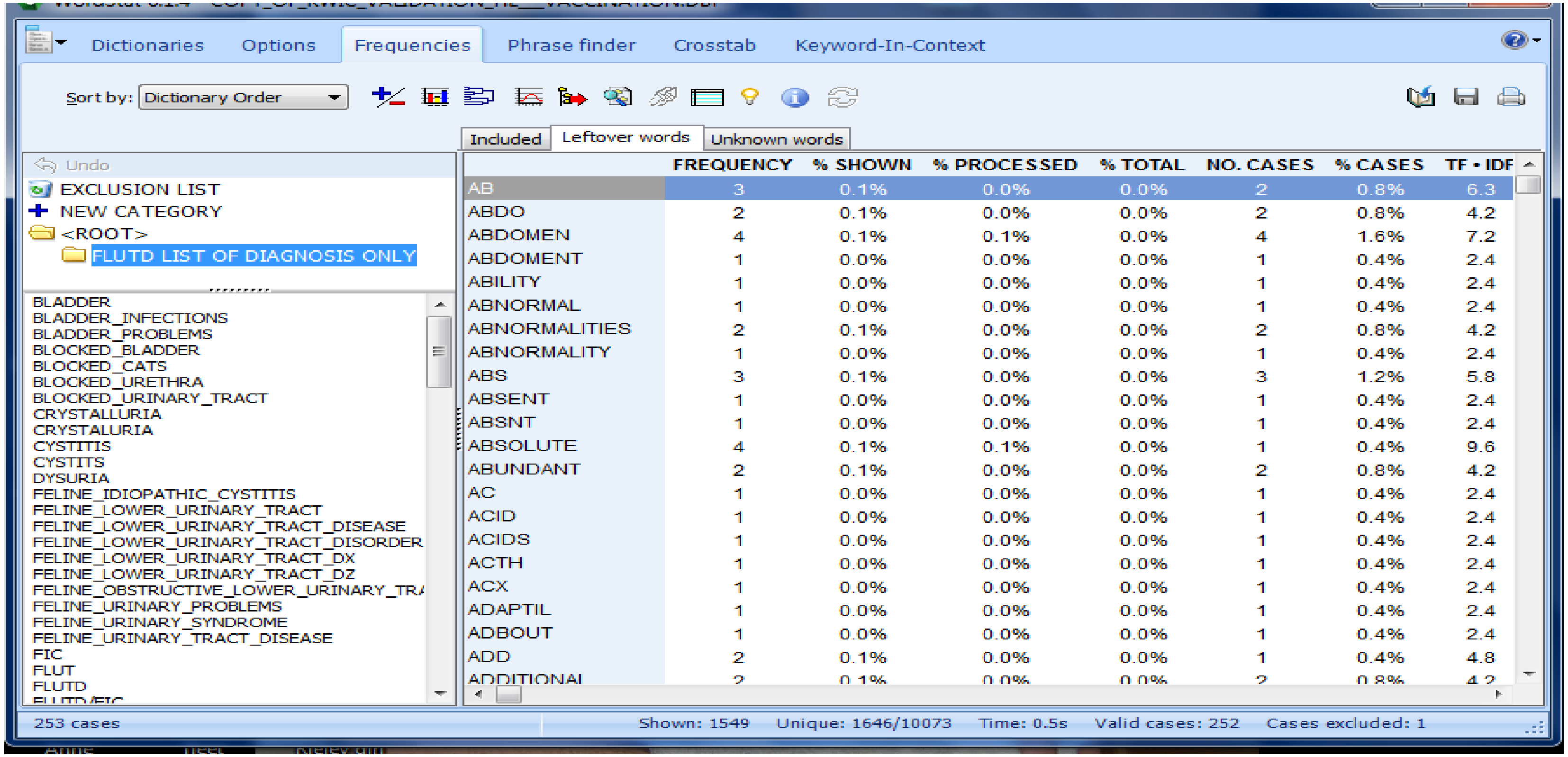


Figure 4-2 Screenshot of WordStat dictionary 1.0 and EPR frequency analysis with leftover word list

4.3 Results

4.3.1 Sample Population

The 8 week extraction produced a total EPR dataset of 2519 patient records from 775 individual animals, 22 species, and were a combination of consultation notes, diagnosis codes, invoice records, test results and reports as previously described (Chapter 2). The 2519 patient records included 645 records relating to feline visits by 198 cats. Removal of the invoice records from the 645 feline records produced a group of non-invoice feline patient records which formed a sample population of 286 records for text analysis (Figure 4-3).

4.3.2 Veterinary Review – The Gold Standard

The veterinarian reviewers identified 4 out of a total 198 feline patients which had reference to clinical signs or diagnosis for FLUTD in their patient record. One feline patient (Animal ID 12410) had two separate records relating to FLUTD (Table 4-2). Therefore the 4 out of 198 (2%) feline patients were associated with 5 out of 286 records (1.75%) referencing clinical signs associated with FLUTD.

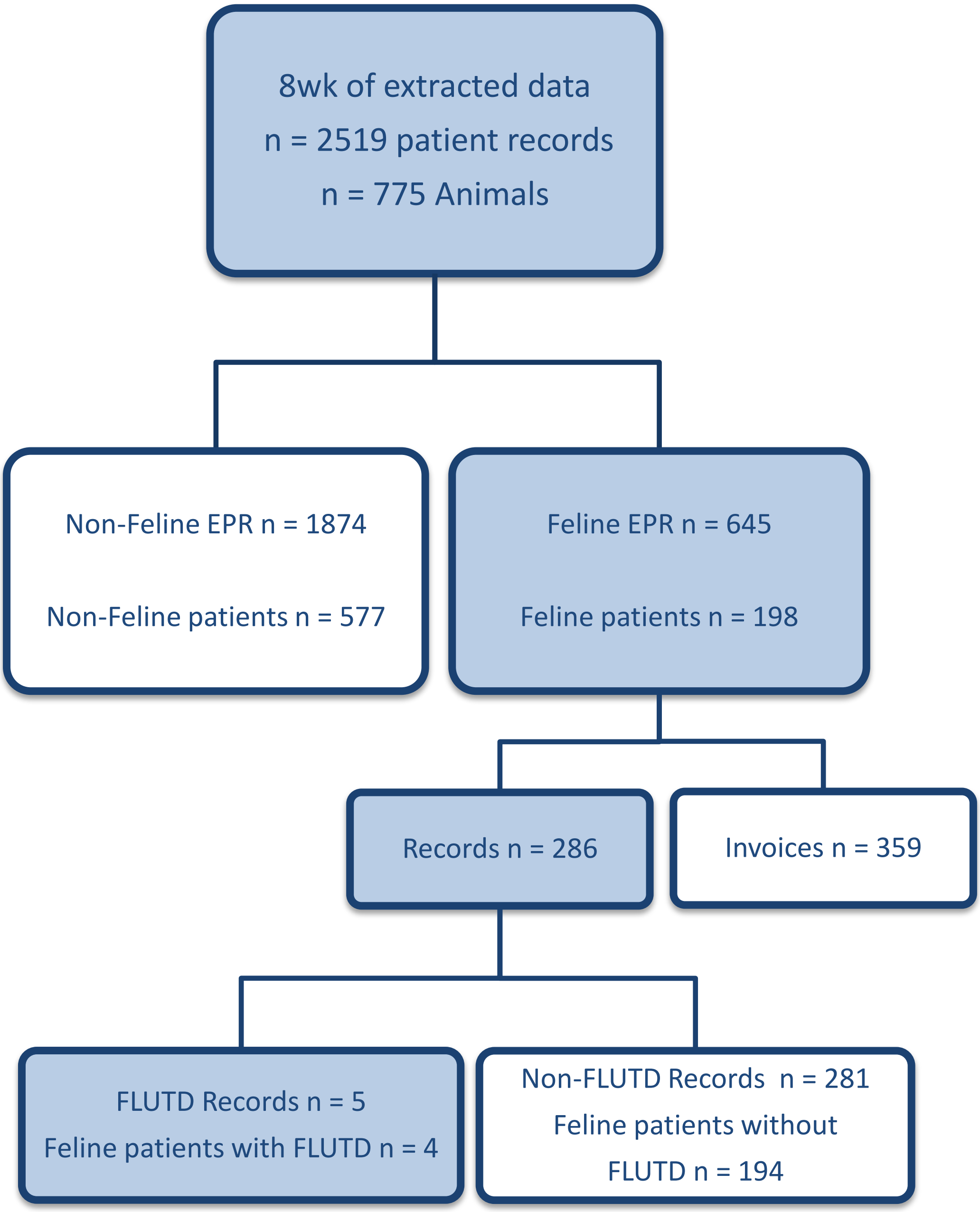


Figure 4-3 Nested target diagram of study population using a 'gold standard' manual search to identify the number of feline patients with FLUTD.

Table 4-2 Table of feline patient records identified by veterinary review to have clinical signs associated with FLUTD.

No of records	Animal ID	Textual information selected by WordStat	Diagnostic code
1	**59**	Showing signs of **cystitis** but now resolved. **Bladder** small but non painful. Adv **urine** sample to check for crystals. Given katkor. Put onto **urinary**.	**Consultation**
2	**5773**	Not eaten much over past week to 10 days. Heart rate 190 bpm. Has been showing signs of **cystitis**. No vomiting. Has dought cranial abdomen. Suspect something going on in stomach or liver. Try b12 nad if no imp a week then adv checking T4 levels.	**Consultation**
3	**12410**	Insurance claim sent to Direct Line for continuation of **Dysuria**.	**Insurance claim**
4	**12410**	**Bladder** scan shows 12-20 1-2mm opacities in **bladder**, v small stones no sign of larger stone notes prev. U+ sample diff to get as **bladder** small pH = 6.5 so ok. Adv try on s/d as will reduce pH further recheck before end of bag 3-5wks.	**Consultation**
5	**13514**	Overnight some blood in **urine** thirst appetite normal thirst appetite normal t38.1 fed Purina low pr + denes **bladder** ok highly strung req us start metacam	**Cystitis**

Dotted line highlights the same animal found twice due to the records containing two diagnostic codes for the same complaint.

4.3.3 The precision of the FLUTD text mining dictionaries.

4.3.3.1 Dictionary 1.0

The survey analysis produced a list of 123 individual phrases or terms relating to cat, FLUTD and bladder from a total of 1096 responses. Therefore the first form of the dictionary, Dictionary 1.0, was composed of 123 phrases or words selected from the survey responses for cat and bladder. Only 5 terms from dictionary 1.0 were found in the feline EPR. These 5 terms occurred many times throughout the extracted EPR (total n = 16) and highlighted the records of 8 cats in total (Table 4-3).

The 5 terms in dictionary 1.0 found in the records were; Bladder, cystitis, urinary, urine, dysuria. All but one of the terms was found within the Text Entry field of the animals EPR. One of the terms was found within the Diagnosis (coded) field of the EPR highlighted in yellow below although this patient also had terms within their Text entry field see details in bold in the selected WordStat result below. The full WordStat output is presented in Appendix 8 (Dictionary 1.0, Case#169, Animal 13514, 12-14).

1. [Case #169 NS1_ANIMAL = 13514 VARIABLE = NS1_TEXTEN]

   overnight some blood in urine thirst appetite normal t38.1 fed purina low pr + denes **bladder** ok highly strung req us start metacam

2. [Case #169 NS1_ANIMAL = 13514 VARIABLE = NS1_DIAGNO]

   Cystitis

3. [Case #169 NS1_ANIMAL = 13514 VARIABLE = NS1_TEXTEN]

   overnight some blood in **urine** thirst appetite normal t38.1 fed purina low pr + denes bladder ok highly strung req us start metacam

Table 4-3 WordStat dictionary search using Dictionary 1.0 to identify cats with terms related to FLUTD in their EPR

	Veterinary Review "Gold Standard"		
Dictionary 1.0	Positive	Negative	Total No cats
Positive	4	4	8
Negative	0	190	190
Total No cats	4	194	198

Sensitivity (*Se*) = 4/4 = 100% (95% CI 51.0% to 100%)

Specificity (*Sp*) = 190/194 = 98% (95% CI 94.8% to 99.2%)

PPV 4/8 = 50% (95% CI 21.5% to 78.0%)

NPV 190/190 = 100% (95% CI 98.0% to 100%)

4.3.3.2 Dictionary 1.1

Analysis of the patient records using Dictionary 1.1 produced a larger list of patient records than Dictionary 1.0 (see Appendix 8 - Dictionary 1.1). Fourteen separate terms from dictionary 1.1 were identified within the EPR. Twenty records, involving 18 feline patients Table 4-4), were identified with one or more terms from dictionary 1.1 present in the clinical notes. The 14 terms identified were; Bladder, Cystitis, Urinary, Urine, Ultrasound, Feliway, Urinating, PUPD, Dysuria, pH, Stone, Straining, Urinalysis, Urinate.

Table 4-4 Evaluation of the WordStat dictionary search of extracted feline consultation data from a sentinel veterinary PMS using Dictionary 1.1.to identify cats with FLUTD.

	Veterinary Review "Gold Standard"		
Dictionary 1.1	Positive	negative	Total No cats
Positive	4	14	18
Negative	0	180	180
Total No cats	4	194	198

Sensitivity (*Se*) = 4/4 = 100% (95% CI 51.0% to 100%)

Specificity (*Sp*) = 180/194 = 93% (95% CI 88.2% to 95.6%)

PPV 4/18 = 22% (95% CI 17.0% to 54.0%)

NPV 180/180 = 100% (95% CI 97.9% to 100%)

4.3.3.3 Dictionary 1.2

A review of the results from the original text mining search using Dictionary 1.0 found that the term ‘bladder’ appeared in a number of records that were not related to FLUTD. Although not part of the original methods, these findings led to the term ‘bladder’ being removed from the original Dictionary 1.0 to create a third Dictionary 1.2 and the analysis was run through the WordStat program again. The results of text mining the feline EPRs using Dictionary 1.2 found 4 terms relating to FLUTD were present in the record; Urine, Dysuria, Cystitis, Urinary. The 4 terms appeared a total of 7 times within the EPRs involving 4 cats’ records. Veterinary review, the gold standard test for this research, found all four cats were genuine FLUTD patients with an FLUTD term within their patient record (

Table 4-5).

One consultation record, which held information relating to FLUTD, was missed due to the removal of the term bladder from the dictionary. However although the consultation was missed the patient was still identified as the term dysuria was used within a second consultation record related to this patient (Table 4-2, Animal ID 12410).

Table 4-5 Evaluation of the WordStat dictionary search of extracted feline consultation data from a sentinel veterinary PMS using Dictionary 1.2.to identify cats with FLUTD.

	Veterinary Review "Gold Standard"		
Dictionary 1.2	Positive	Negative	Total No cats
Positive	4	0	4
Negative	0	194	194
Total No cats	4	194	198

Sensitivity (*Se*) = 4/4 = 100% (95% CI 51.0% to 100%)

Specificity (*Sp*) = 194/194 = 100% (95% CI 98.6% to 100%)

PPV 4/4 = 100% (95% CI 51.0% to 100%)

NPV 194/194 = 100% (95% CI 98.6% to 100%)

4.3.3.4 Statistical Measures

The dictionaries produced very different search results as presented below in Table 4-6. Dictionary 1.2 was found to be the most sensitive. The least sensitive was Dictionary 1.1 which contained the largest number of search terms. However the increase in the number of results and decrease in sensitivity for Dictionary 1.1 was primarily due to the terms feliway, pH, urinalysis and PUPD, which identified a high number of non-FLUTD cases.

The results of all three dictionaries (v1.0, v1.1 & v1.2) were compared as presented below in Table 4-6.

Table 4-6 Search results from all three text mining dictionaries

Dictionary version	Total N^{o} of terms found in record	N^{o} of individual terms	N^{o} records with a term present	N^{o} of cats identified with ref to FLUTD in their record	N^{o} of true positive cats with FLUTD*
v. 1.0	16	5	9	8	4/8
v. 1.1	33	14	20	18	4/18
v. 1.2	7	4	4	4	4/4

*All true positive FLUTD cases present within the total number of cat records identified by the search methodology for all three versions of the dictionary.

4.4 Discussion

Dictionaries are designed by attempting to list the structure and terminology of a human language, organised alphabetically (Hersh, 2003b). In text mining it is necessary to have some knowledge of the terminology used in the text to be able to successfully identify the information of interest (Szolovits, 2003).

Despite a limited knowledge of the exact veterinary medical terminology for FLUTD used by the sentinel practice, the clinical terms dictionary designed for this study using the Wordstat Dictionary text mining function proved to be an excellent tool for search and retrieval of FLUTD patient records. This text mining program also offered a feedback function for leftover words which allowed the dictionaries to be further developed. In addition the trial provided an excellent insight into the terminology used by vets in this sentinel practice to record clinical conditions or signs of disease.

The initial search of the records using Dictionary 1.0 produced an excellent sensitivity (100%) and specificity (98%) but a low positive predictive value at only 50%. Negative predictive values were calculated for all the dictionaries and were found to be 100%. Whilst this is also an excellent result suggesting the method can reliably identify animals that are disease free, in the case of this research the identification of diseased animals from the sample population was the priority. The low positive predictive value was a consequence of the high number of results which were not true cases of

FLUTD because the search term ‘bladder’ appeared many times within the patient record increasing the number of false positives (reduced specificity); this may be because the bladder is a common organ to examine during any routine health check and therefore may be written frequently within the EPR. As such it should be considered a non-specific term which would not be recommended for use in search dictionaries.

The presence of the term bladder in the first two dictionaries meant a high number of results did not have terms relating purely to FLUTD in the patients records. However the true positives were still found among the false positives hence the high sensitivity of this methodology but slightly lower specificity and positive predictive power. To identify all the true positive animals from the whole dataset it is desirable for the results of the search to have a high sensitivity and PPV. Specificity is an important factor in this equation, therefore although high sensitivity is desirable the specificity should also be maintained at a relatively high level for the PPV and subsequently the usefulness of the method to remain high which is often a challenge as with measures of sensitivity and specificity one compromises the other (Altman and Bland, 1994).

The small number of cases identified resulted in wide confidence intervals and therefore less certainty. Dictionary 1.0 produced only a small number of results, finding only 8 cats in total, four of which (50%) were true FLUTD cases. The follow-up check, to identify which of the 8 records were true positive cases, therefore took very little time. However in a real situation using much larger EPRs and where the search may produce lists of thousands of results, this lack of precision would double the work for follow-up review and be impractical. Therefore it may be necessary to accept false positives will occur in a search where 100% sensitivity is desired.

For the dataset examined here, on only one occasion did veterinarians record a diagnosis (Cystitis) within the diagnosis field of the patient record. This highlights an important trend in the data which may suggest the diagnosis codes are only used very rarely to record a clinical condition, certainly in this example. This is an important finding and highlights the limitations for the success of search and retrieval methods using only the coded portion of an EPR, which is a common finding in text mining of medical records (Stein et al., 2000, Hersh, 2003b).

It may be assumed the more terms used to search the more likely it is that cases of disease would be identified (Petrie and Watson, 2006, Petrie and Sabin, 2009). However this would only be the case if the terms added are appropriate, specific and used by practitioners. The addition of more terms from the left-over words in WordStat frequency analysis to create Dictionary v1.1 actually reduced the specificity of the search, and as a consequence the PPV was reduced. This drop is most likely because the terms added were non-specific terms.

The estimation of terms likely to be used, and therefore the basis of dictionary 1.0, was a survey of the profession and predominantly consisted of common conditions or complaints rather than clinical signs. As a result, many of the terms used for this dictionary search were not specific to FLUTD or present in the patient record and only a small proportion (4%) of all terms used within the dictionaries actually appeared within the notes.

The objective of this study was method development and it was only posible to estimate which terms were more specific once the searches were complete. Dictionary v1.0 included many terms and allowed a full review of the EPR. Once the larger search had been performed it was possible to use the content analysis software to refine the method. During this process it became apparent the term bladder was not specific enough for inclusion. By removal of the term 'bladder' the final dictionary version 1.2 produced results where sensitivity remained at 100% but the specificity, positive and negative predictive values were then also found to be excellent with all at 100%.

The most successful list was therefore Dictionary v1.2. The list had very few terms which could be found in the EPR, however the number of positive cases identified was 100%. This therefore was due to the more disease-specific terms selected for the dictionary and the removal of very non-specific terms, i.e. bladder. This result is the ideal. As there is currently no standard veterinary terminology, it is important to identify the few specific terms used by the majority of veterinarians for decribing clinical signs of disease to create clinical terms dictionaries with disease specific terms and a high level of positive predictive value.

To identify all true positive cases and true negative cases of one disease within a veterinary electronic patient record is a difficult task. The task is further complicated by the prevalence of any particular disease within the population and the need to predict the disease specific search terms required to find those positive cases and to exclude the negative cases. A limitation of content analysis of the veterinary EPR is the lack of a standard terminology used by the veterinary profession. Clinical codes are used commonly in human practice-based research to measure the prevalence and incidence of disease along with free text, where confidentiality allows, and prescription information. However research within the human healthcare field using data extracted from general practice databases have reported similar frustrations caused by varied and random terminology (Lawrenson et al., 1999).

In a review published in the Journal of Public Health Medicine, Lawrenson et al., (1999) reports the value of the prescription file of general practice databases as the most useful attribute to practice-based research of this type. In GP practice the prescription record is usually complete, often containing up to 95% of all perscriptions issued by a GP. Details of drug, dose and duration are all recorded. However much of the information relating to how to take the drug is contained within the free text portion of the record and may be written in many different ways by each GP (2 x daily, B.D, twice a day).

Lawrensen uses this example to illustrate the frustration of having all the information one needs to understand the patient data, but not the means with which to know all terms or where all the information is recorded to convert them into one simple dataset for analysis. Text mining methods using a dictionary of these instructions may be used to select out the narrative associated with prescribing.

The review by Lawrenson et al., (1999) offers an interesting next step for this research. The results from Laurensons study suggests there may be some value in examining the invoice data, to identify the frequency of different treatments used which may add to the richness of the extracted data for analysis. The invoice data was not included for this study due to the repetition of information. However assuming the duplication can be controlled for there may be value in searching the invoice section of the EPR to allow further investigation of prescribing behaviour in first opinion veterinary practice and also a comparison of how well the invoice section and consultation records are matched.

4.5 Conclusion

Practice-based research within the veterinary profession would benefit from the standardisation of veterinary terminology; unfortunately this is not currently in place and may not be for some time, if ever, and to try to predict the terms used by veterinarians prior to text mining is impractical.

The method described here offers a quick and efficient means to identify a high proportion of cases in a relatively short timeframe and provides an excellent solution to the current difficulties for practice-based research. Until such time that a veterinary standardised terminology exists the use of clinical terms dictionary searching will be invaluable as a tool for practice-based research.

Chapter Five

5 Data capture and precision within the electronic patient record.

5.1 Introduction

The introduction of computers to veterinary practice has streamlined the consultation process by assisting the process of billing, ordering, stock control, sample analysis and the recording of patient information (Hersh, 2003b). Once introduced, the progression from paper to computerised practice records was rapid and appears to have been accepted readily by the practicing veterinarian (McCurdy, 2001). However, the conversion has primarily been to support the clinical care of patients, with more efficient record keeping, and the management of the practice, by assisting with stock control and billing, not necessarily research and although the two are not mutually exclusive, research is somewhat of a by-product of the computerisation of veterinary practice.

In the human field, an improvement in record keeping has been proposed as one method of improving patient care (Stott and Davis, 1979, Harman et al., 2012) and the introduction of computerised practice management software (PMS) systems and electronic patient records (EPR) has been found to offer an improvement in the organisation and efficiency of record keeping and therefore patient care in both human (Blumenthal and Tavenner, 2010) and veterinary medicine (Lam et al., 2007a). Although research to support this has been criticised as inconsistent (Harman et al., 2012)

Practice management software systems were designed primarily as only an accountancy system to assist with billing and stock control (van Bemmel and Musen, 1997, Shortcliffe and Blois, 2003) and has only recently been utilised as a tool for research (McCurdy, 2001, Faunt et al., 2007). It is therefore understandable that without this as an original design consideration it is unlikely to be a perfect instrument for research.

While PMS systems and veterinary EPRs may not be perfectly designed for research, associated and emerging computer-based technologies may help with data accessibility. Methodologies such as Natural language processing (NLP), Information extraction (IE) Content analysis (CA) and Text mining (TM) should assist to a great extent with data extraction, management and analysis (Vandeweerd et al., 2012a, Vandeweerd et al., 2012b).

The success of veterinary practice-based research, using information captured within the EPR, may therefore depend less on accessibility to the information and more on the accuracy and quality of data recorded during the veterinary consultation (Bernstein et al., 1993). Limited or incomplete information recorded in the electronic patient record of veterinary PMS systems would be a major hurdle for many studies using veterinary informatics. It is therefore important to understand how accurately an EPR reflects the content of the actual consultation, only then can any limitations of electronically collected data be fully realised. Examining consultations directly and then extracting the records made during those same consultations would enable the quality and quantity of data captured in the EPR during veterinary consultation to be compared. Any disparity between information recorded in the EPR to that discussed during the consult would be identified for further investigation.

The purpose of this research was to examine the information captured and extracted from the practice PMS system from a working veterinary practice EPR during one week of consultations and compare with that collected by direct observation by a qualified veterinary researcher during the consultation.

5.2 Materials and Methods

5.2.1 Data Collection

Data collection encompassed one week of consultations by the practice veterinarians (n=2) at a sentinel veterinary practice (Chapter 2). Data was recorded from 14^{th} - 18^{th} May 2012 between the hours of 8.30am – 8pm.

5.2.1.1 Collection of consultation data by XML schema extraction

The data recorded in the EPR of the sentinel PMS system during the week was extracted by the researcher (JJD) using the integration of the Clinical Evidence XML schema into the sentinel practice PMS as previously described (Chapter 2). All consultations recorded during one week were extracted by the sentinel senior veterinarian and forwarded to JJD as an XML data file for analysis.

5.2.1.2 Collection of consultation data by direct observation

Direct observation, with prior agreement, was made of a number of consultations (n = 36) performed by 2 veterinarians at the sentinel practice from 14^{th} – 18^{th} May 2012 by a qualified veterinarian (NR). The data was collected as part of a separate, ethically approved, observational study. NR recorded details of the consultation using a standardised data collection form (Figure 5-1) for the purpose of real time observational data collection and analysis for use in first opinion veterinary practice (*unpublished data,* N. Robinson *Ph.D Thesis 2014*).

During the observation, NR recorded the date and time of the consultation and the problems presented by the client to the veterinarian including the order in which the problems were discussed. Data collected included signalment of the animals visiting the practice, the reason for the visit and any additional problems which were discussed. In addition treatments prescribed or outcomes of previous treatment and any medical or treatment decisions made were recorded. The data was collected during the consultation onto the data collection form by hand which was then scanned into a database using the Teleform program[15].

The consultations observed and recorded by NR were matched to the record collected via EPR extraction by JJD. All data was anonymised, therefore observed and extracted consultations were matched by JJD using date and time of consultation and matching signalment data such as breed, age, sex and neuter status.

[15] http://www.cardiff-teleform.com

Date (DD/MM/YY) ☐☐ / ☐☐ / ☐☐ | Practice ☐☐☐ | Consult. No. ☐☐☐☐ | Animal. No. ☐☐☐☐ / ☐ | Vet Initials ☐☐☐

VN Client type ☐ Private ☐ PDSA

Questionnaire

1. Were multiple animals presented? Complete a separate questionnaire for each animal. ☐ Yes ☐ No
2. Select the best description of the type of case from the following options:
   ☐ First Consult ☐ Recheck ☐ Elective Euth ☐ Recurrent ☐ 2nd Op
   ☐ Ongoing: Acute ☐ Ongoing: Chronic ☐ Monitoring ☐ Prev Med ☐ Admit/Discharge
   ☐ Other
3. Which species was presented during the consult?
   ☐ Dog ☐ Cat ☐ Rabbit ☐ Ferret
   ☐ Rodent ____ ☐ Bird ____ ☐ Reptile ____ ☐ Other ____
4. What was the animals breed?
   Records: ____ Vet: ____ Owner: ____
5. What was the animals age?
   Records: ☐☐ Y ☐☐ M ☐ W ☐ D Vet: ☐☐ Y ☐☐ M ☐ W ☐ D Owner: ☐☐ Y ☐☐ M ☐ W ☐ D
6. What was the animals sex including neutering status?

	MN	ME	FN	FE	MU	FU	U
Records:	☐	☐	☐	☐	☐	☐	☐
Vet:	☐	☐	☐	☐	☐	☐	☐
Owner:	☐	☐	☐	☐	☐	☐	☐

7. Was a clinical exam performed? ☐ Yes: full exam ☐ Yes: focused exam ☐ No
   If yes, were any abnormalities detected? ☐ Yes ☐ No
8. Was the animal weighed during the consultation period? ☐ Yes ☐ No

Page1/3

	Problem 1	Problem 2	Problem 3	Problem 4
Problem summary/ clinical signs				
Related C.E. findings?	☐ Yes ☐ No ☐ N/A	☐ Yes ☐ No ☐ N/A	☐ Yes ☐ No ☐ N/A	☐ Yes ☐ No ☐ N/A
Raised by	☐ Owner ☐ Vet ☐ Prompt	☐ Owner ☐ Vet ☐ Prompt	☐ Owner ☐ Vet ☐ Prompt	☐ Owner ☐ Vet ☐ Prompt
Bodysystem affected	☐ Skin ☐ MSK ☐ Neuro ☐ Eyes ☐ Urin ☐ Renal ☐ Repro ☐ GI ☐ Cardio ☐ Haemo ☐ Resp ☐ Endo ☐ Dental ☐ Non-sp ☐ Prev Med ☐ Behav	☐ Skin ☐ MSK ☐ Neuro ☐ Eyes ☐ Urin ☐ Renal ☐ Repro ☐ GI ☐ Cardio ☐ Haemo ☐ Resp ☐ Endo ☐ Dental ☐ Non-sp ☐ Prev Med ☐ Behav	☐ Skin ☐ MSK ☐ Neuro ☐ Eyes ☐ Urin ☐ Renal ☐ Repro ☐ GI ☐ Cardio ☐ Haemo ☐ Resp ☐ Endo ☐ Dental ☐ Non-sp ☐ Prev Med ☐ Behav	☐ Skin ☐ MSK ☐ Neuro ☐ Eyes ☐ Urin ☐ Renal ☐ Repro ☐ GI ☐ Cardio ☐ Haemo ☐ Resp ☐ Endo ☐ Dental ☐ Non-sp ☐ Prev Med ☐ Behav
Diagnostic tests	☐ In-cons ☐ Post-cons ☐ None	☐ In-cons ☐ Post-cons ☐ None	☐ In-cons ☐ Post-cons ☐ None	☐ In-cons ☐ Post-cons ☐ None
In Cons				
Post Cons				
Diagnosis	☐ Open ☐ Definitive ☐ Presumed ☐ Prev. Dx. ☐ Working ☐ N/A	☐ Open ☐ Definitive ☐ Presumed ☐ Prev. Dx. ☐ Working ☐ N/A	☐ Open ☐ Definitive ☐ Presumed ☐ Prev. Dx. ☐ Working ☐ N/A	☐ Open ☐ Definitive ☐ Presumed ☐ Prev. Dx. ☐ Working ☐ N/A
Outcome	☐ Nothing ☐ Manage ☐ Work up ☐ Ther. Tx ☐ Euth ☐ Prop. Tx ☐ Refer ☐ Other	☐ Nothing ☐ Manage ☐ Work up ☐ Ther. Tx ☐ Euth ☐ Prop. Tx ☐ Refer ☐ Other	☐ Nothing ☐ Manage ☐ Work up ☐ Ther. Tx ☐ Euth ☐ Prop. Tx ☐ Refer ☐ Other	☐ Nothing ☐ Manage ☐ Work up ☐ Ther. Tx ☐ Euth ☐ Prop. Tx ☐ Refer ☐ Other

Page 2/3

☐☐ / ☐☐ / ☐☐ Date (DD/MM/YY)   ☐☐ Practice   ☐☐☐☐ / ☐ Animal. No.

	Problem 5	Problem 6	Problem 7	Problem 8
Problem summary/ clinical signs				
Related C.E. findings?	☐ Yes ☐ No ☐ N/A	☐ Yes ☐ No ☐ N/A	☐ Yes ☐ No ☐ N/A	☐ Yes ☐ No ☐ N/A
Raised by	☐ Owner ☐ Vet ☐ Prompt	☐ Owner ☐ Vet ☐ Prompt	☐ Owner ☐ Vet ☐ Prompt	☐ Owner ☐ Vet ☐ Prompt
Bodysystem affected	☐ Skin ☐ MSK ☐ Neuro ☐ Eyes ☐ Urin ☐ Renal ☐ Repro ☐ GI ☐ Cardio ☐ Haemo ☐ Resp ☐ Endo ☐ Dental ☐ Non-sp ☐ Prev Med ☐ Behav	☐ Skin ☐ MSK ☐ Neuro ☐ Eyes ☐ Urin ☐ Renal ☐ Repro ☐ GI ☐ Cardio ☐ Haemo ☐ Resp ☐ Endo ☐ Dental ☐ Non-sp ☐ Prev Med ☐ Behav	☐ Skin ☐ MSK ☐ Neuro ☐ Eyes ☐ Urin ☐ Renal ☐ Repro ☐ GI ☐ Cardio ☐ Haemo ☐ Resp ☐ Endo ☐ Dental ☐ Non-sp ☐ Prev Med ☐ Behav	☐ Skin ☐ MSK ☐ Neuro ☐ Eyes ☐ Urin ☐ Renal ☐ Repro ☐ GI ☐ Cardio ☐ Haemo ☐ Resp ☐ Endo ☐ Dental ☐ Non-sp ☐ Prev Med ☐ Behav
Diagnostic tests	☐ In-cons ☐ Post-cons ☐ None	☐ In-cons ☐ Post-cons ☐ None	☐ In-cons ☐ Post-cons ☐ None	☐ In-cons ☐ Post-cons ☐ None
In Cons				
Post Cons				
Diagnosis	☐ Open ☐ Definitive ☐ Presumed ☐ Prev. Dx. ☐ Working ☐ N/A	☐ Open ☐ Definitive ☐ Presumed ☐ Prev. Dx. ☐ Working ☐ N/A	☐ Open ☐ Definitive ☐ Presumed ☐ Prev. Dx. ☐ Working ☐ N/A	☐ Open ☐ Definitive ☐ Presumed ☐ Prev. Dx. ☐ Working ☐ N/A
Outcome	☐ Nothing ☐ Manage ☐ Work up ☐ Ther. Tx ☐ Euth ☐ Prop. Tx ☐ Refer ☐ Other	☐ Nothing ☐ Manage ☐ Work up ☐ Ther. Tx ☐ Euth ☐ Prop. Tx ☐ Refer ☐ Other	☐ Nothing ☐ Manage ☐ Work up ☐ Ther. Tx ☐ Euth ☐ Prop. Tx ☐ Refer ☐ Other	☐ Nothing ☐ Manage ☐ Work up ☐ Ther. Tx ☐ Euth ☐ Prop. Tx ☐ Refer ☐ Other

Page 3/3

Figure 5-1 Example of a data collection sheet used to record data during the observation.

5.2.2 Data Formatting

The observational data was recorded by hand onto a data collection form whereas the information recorded by the veterinarian at the sentinel practice was entered directly into the EPR. It was assumed the observational record would produce a richer dataset as all events, discussions and treatments would be logged, regardless of outcome or relevance, by the observer. Therefore the information recorded by observation was used as the gold standard with which to compare the EPR extracted information.

The information recorded by the practice veterinarian during the consultations into the practice computer was extracted electronically and then transferred by hand onto a paper data collection form to allow the datasets to be compared directly as two identical formats. To ensure the extracted data was transferred accurately and appropriately onto the data collection form, it was performed in duplicate by two veterinary reviewers (NR and RD) from the Centre for Evidence-based Veterinary medicine, NR was the original observer and RD an additional independent veterinarian, Figure 5-2 presents an outline of the process of data collection, aggregation and analysis.

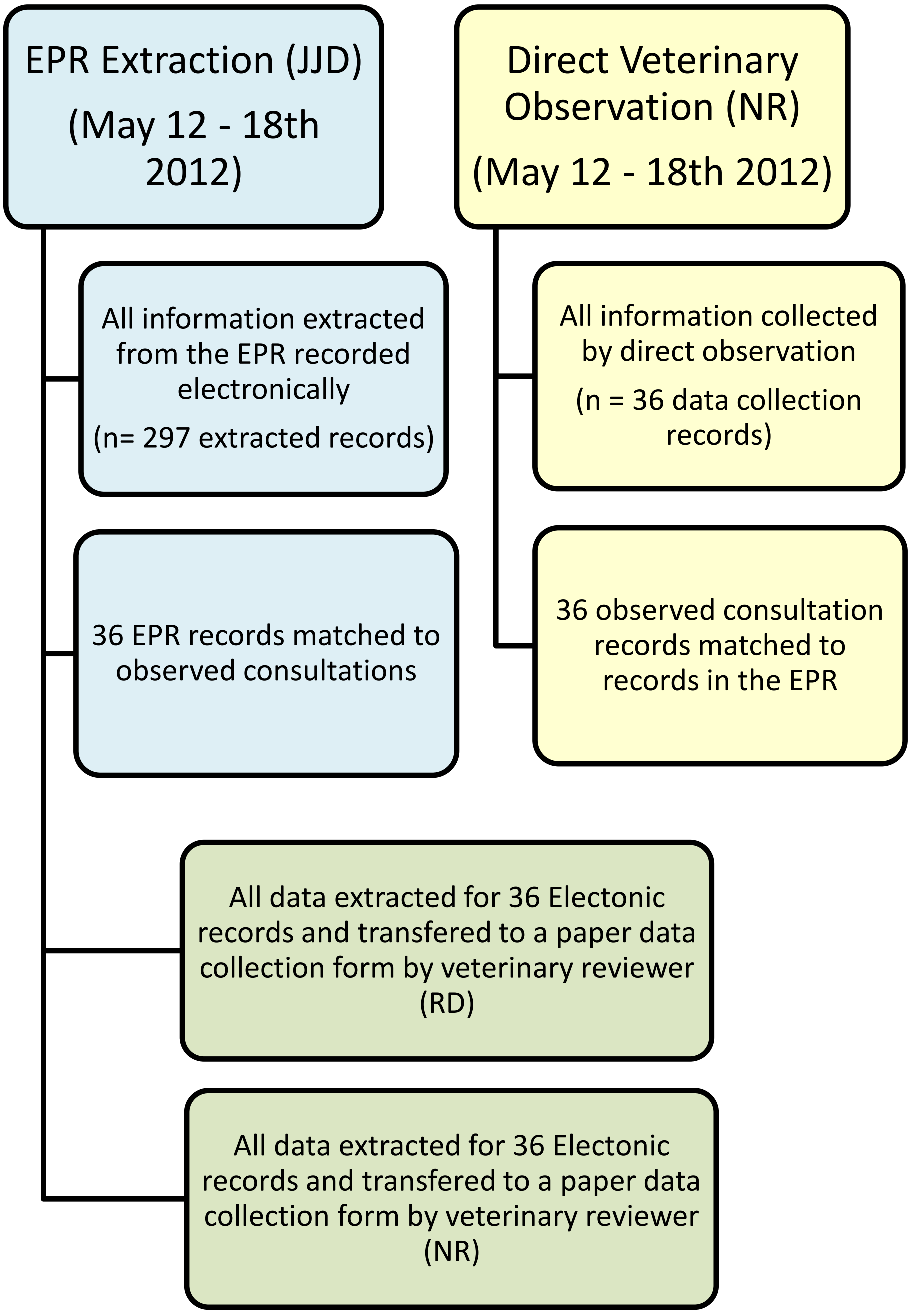


Figure 5-2 Flow diagram describing the process of data collection, aggregation and analysis.

5.2.3 Data Analysis

5.2.3.1 Study 1 – Measure of Agreement

All data extracted from the PMS system were imported into a Microsoft Excel 2010 Workbook. Data analysis initially involved an assessment of the data collected by both methods (extraction and observation) and transfer of the extracted data onto a paper data collection form for comparison. Two reviewers (NR & RD) performed the assessment and information transfer for each of the 36 animal consultations. A measure of agreement between the reviewers for their interpretation of events during the consultation as recorded in the EPR and subsequent transfer of this information onto a paper data collection form was performed using Cohen's Kappa analysis index of inter-rater reliability.

Statistical analysis was performed using Minitab 16 (v16.2.2). Cohen's Kappa analysis index of inter-rater reliability was performed following the methods described by Petrie and Sabin (2009). A kappa score between 0.6 – 0.8 suggests a substantial level of agreement according to tables produced by (Landis and Koch, 1977).

5.2.3.2 Study 2 – Comparison of Data Quantity

There was a good level of agreement between the veterinary reviewers for the transfer of information between the two recording formats (EPR and paper form) as determined by Kappa analysis. However veterinary reviewer NRs dataset was chosen for the comparative work as it was the most accurate when compared to the observation data. Therefore the observational dataset was compared to the extracted EPR dataset as reviewed and transferred into paper format by NR.

To investigate differences in quantity and quality of the captured data, using the two methods of recording (extraction and direct observation), basic data was analysed and descriptive statistics calculated including;

1. Number of patient visits recorded by both methods, direct observation and extraction.
2. Number and type of problems recorded for each animal visit by both methods (all problems discussed by vet and/or client)
3. Comparison of consultation types as defined by NR recorded by both methods (e.g. Post op check, vaccination, behaviour problem)

An analysis was made of the number of problems recorded for each consult. Direct comparison of information quantity captured for each consultation by both methods of data collection (observation and EPR extraction) was analysed using sensitivity and specificity analysis of problem number, with direct observation data as the gold standard measure.

The PMS system and extraction method, for the collation of clinical data from first opinion veterinary practice, was compared to the gold standard method of direct observation by calculating sensitivity and specificity analysis scores. The sensitivity and the specificity were calculated. Analysis was based on methods described by Petrie and Watson (2006) and presented in Appendix 6.

5.2.3.3 Study 3 – Comparison of Data Quality

This study investigated veterinary recording and also the effect of consult outcome. By using the anonymous vet code as a vet identifier it was possible to calculate the proportion of problems recorded for each of the 2 veterinarians in the EPR and compare this to the number of problems recorded by NR during the observation for each veterinarian. Also it was possible to examine the type of problem raised and order in which problems were discussed and if this influenced the recording of the information in the EPR. Finally NR described 8 different outcome types on the observation forms, detailed below. A frequency analysis was therefore possible to calculate observed outcome measure against problem recording in the EPR. Consultation outcome recorded for each problem discussed and as defined by NR (in brackets) during observation;

1. Therapeutic TX (Therapeutic treatment)
2. Prophylactic TX (Prophylactic treatment)
3. Nothing (No action taken)
4. Manage (No change in treatment, continue to manage an ongoing condition)

5. Work up (Sample analysis, diagnostic tests)
6. Euthanasia (Euthanasia)
7. Other (None of the above)

Therefore the analysis focussed on the following,

i. The effect of veterinarian recording of information (Vet 1 and Vet 2) was examined.
ii. The order in which problems were presented by the client to the veterinarian during the consult and the type of problem discussed, as defined by NR during the observation.
iii. Finally the EPR records were examined to see if there was a relationship between the outcome of a consultation discussion and the recording of information in the EPR.

5.3 Results

5.3.1 Study 1 – Measure of Agreement

The level of agreement between reviewers NR & RD, for interpretation and subsequent transfer of the extracted data from the EPR to the paper form using Cohen's Kappa analysis was found to be very good, $K = 0.723 \pm 0.12$ SE ($P < 0.0001$). Details of the analysis are presented in Appendix 9.

For the 36 animal visits observed; 99 problems were recorded in total with a mean of 2.75 per animal visit (min 1 – max 5). In total NR reviewed the 36 consults in the EPR and identified a maximum of 64 problems recorded with a mean of 1.78 (range 1-4) problems recorded per consult. RD reviewed the same 36 consults and identified 56 problems recorded in the EPR with a mean of 1.56 (range 1-4) problems recorded per consult, The spread of the data is presented in the graph below (Figure 5-3) showing all data recorded by observation and veterinary review.

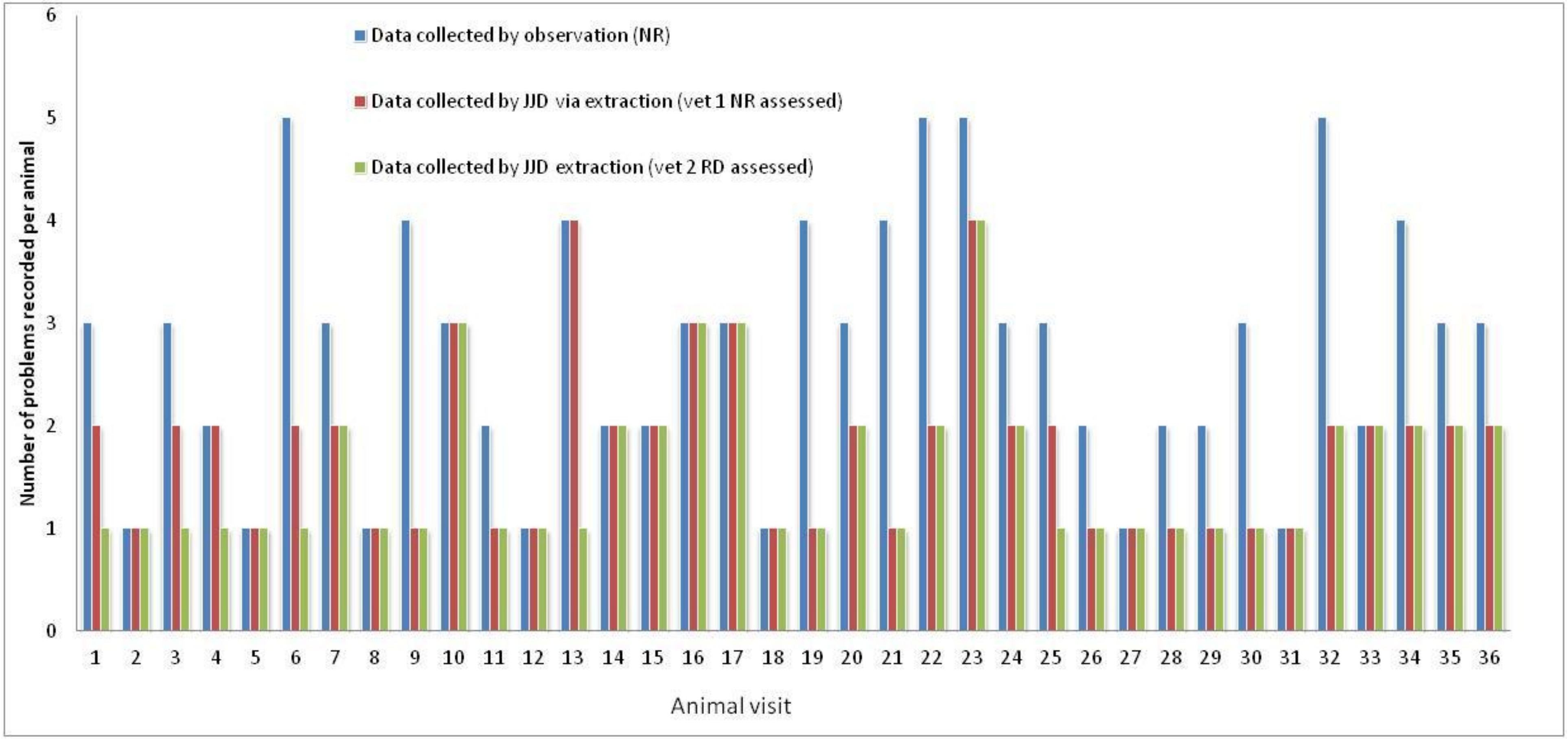


Figure 5-3 Number of problems discussed per consult. Data presented for two methods of data collection, a) direct observation and b) EPR extraction. Extraction was reviewed by two veterinarians (i)NR and (ii)RD.

5.3.2 Study 2 – Comparison of Data Quantity

5.3.2.1 Number of Patient Visits

The extraction of data from the sentinel PMS system for all consultations over the week produced a dataset with 262 visits recorded in the extracted EPR by 158 individual animals. There were 36 animal visits observed directly and recorded by NR (13.7% of all consultations that week). The EPR was found to be complete with all 36 animal visits observed by NR identified and matched to records in the extracted EPR data.

The records collected by both methods, observation and extraction, were compared. The data recorded in the EPRs by the veterinarians was found to represent approximately 65% of all problems discussed when observed directly by NR (Sensitivity = 0.71.9%, 95% CI 61.8% to 80.1%), Table 5-1.

Table 5-1 Number of problems identified by each method (observation vs extraction)

	Direct Observation of Consultations (Gold Standard)		
Extraction of data	Positive	Negative*	Total
Positive	64	0	64
Negative	25	0	25
Total	99	0	99

Sensitivity = 64/99 x 100 = 0.71.9% (95% CI 61.8% to 80.1%) (Specificity (*Sp*) = N/A. (*Specificity could not be measured as the gold standard of direct observation could not observe a consult which didn't occur.)

5.3.2.2 Number of Problems Recorded

Direct observation captured 99 problems discussed in total during the 36 animal consults. However when the 36 extracted records from the EPR were assessed, only a total of 64 problems had been recorded into the EPR by the veterinarian.

The graph in Figure 5-4 presents the differences in recording in relation to problem number recorded by observation compared to that recorded in the EPR. The graph suggests where only one problem was discussed this was recorded in the EPR more consistently than where 2 or more problems were discussed.

The type of problem raised and the order in which the problems were raised by the client for discussion was also examined for each individual consult and cross referenced against the recording of the problem in the EPR. There was no relationship found between the type of problem raised by the client or order of problem raised and the veterinarian’s decision to add the details to the EPR (Appendix 10).

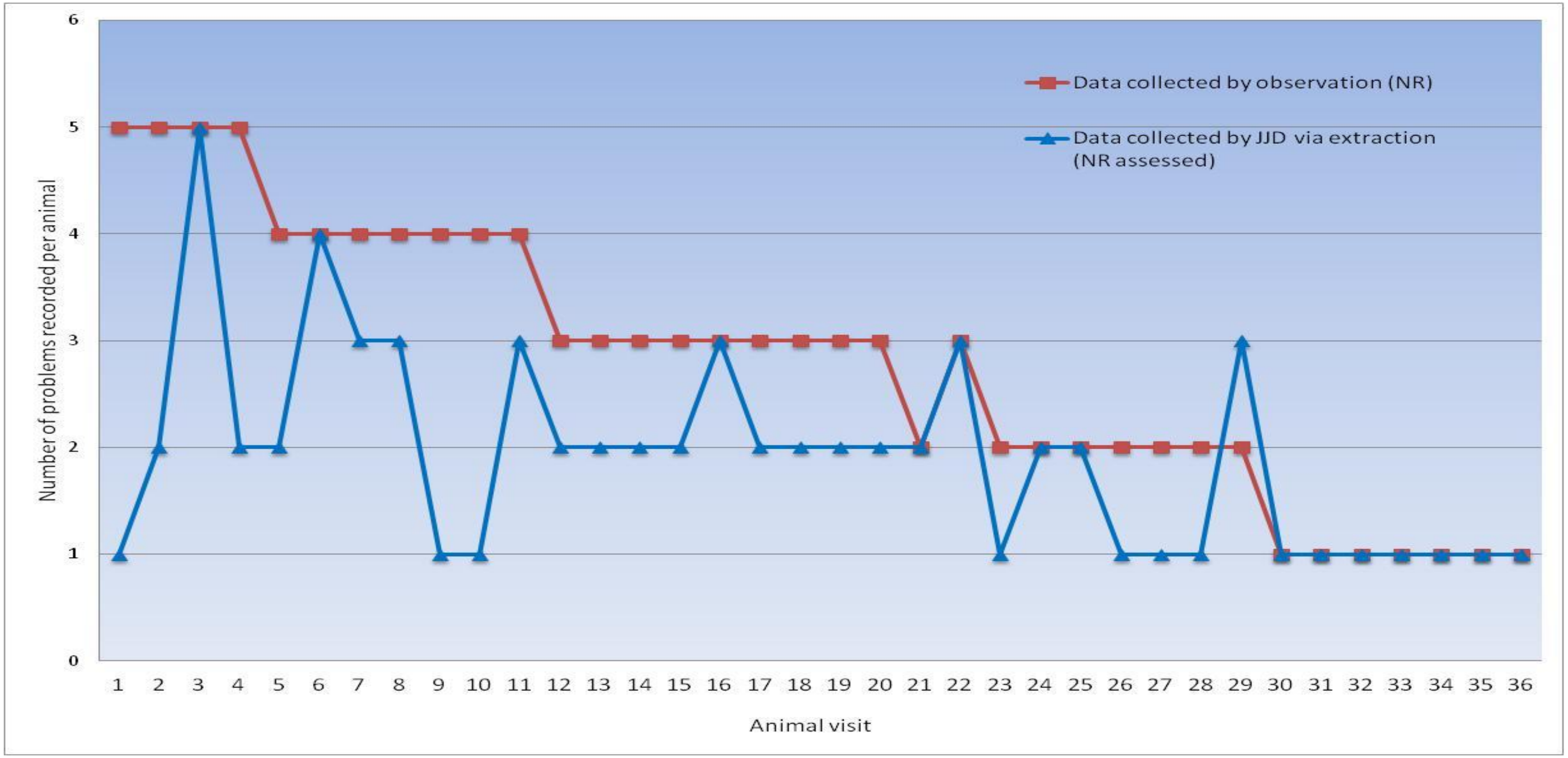


Figure 5-4 Plot to demonstrate the differences between problem number and recording method (n = 36 consults, n = 99 problems, min 1 max 5 per consult)

5.3.3 Study 3 – Comparison of Data Quality

5.3.3.1 Effect of Veterinarian

To assess the effect of veterinarian recording behaviour the individual consultations were separated by vet and compared, Figure 5-5.

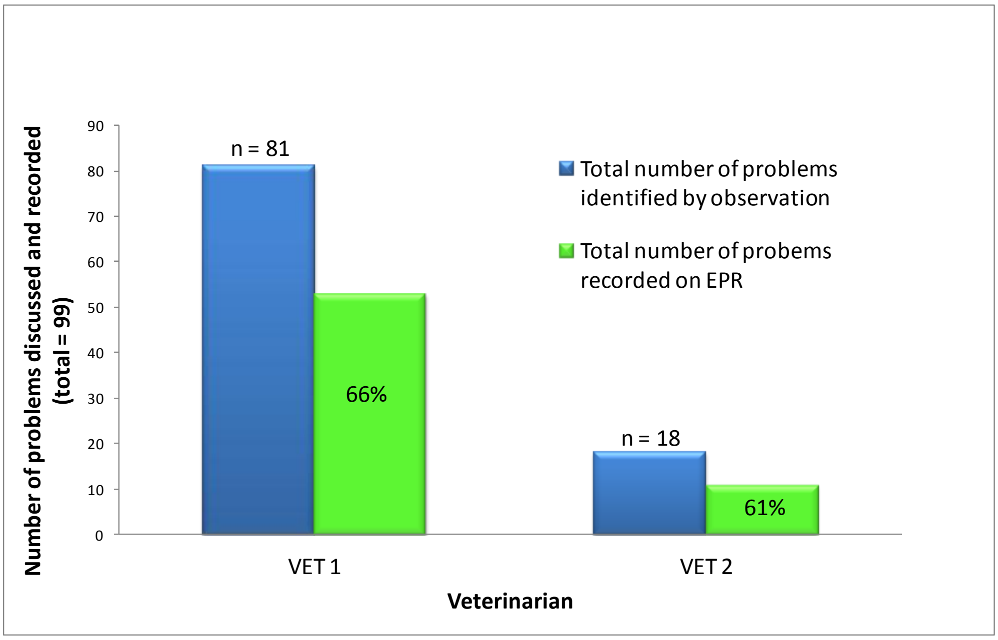


Figure 5-5 Assessment of veterinary recording behaviour in the EPR, comparison of case load and proportion recording of patient data.

Vet 1 performed the majority of the consultations observed and as a consequence addressed a high number of the problems presented by clients *(Vet 1, 30 consultations, n = 81 problems examined)* whereas Vet 2 performed very few *(Vet 2, 6 consultations, n = 18 problems examined).* However, the recording behaviour between the two vets was found to be comparable with both vets recording a similar proportion of problems in the electronic record, (*Vet 1 66% of problems recorded on the EPR, Vet 2 61% of problems recorded on the EPR,* Appendix 11).

5.3.3.2 Effect of Outcome

As the records were examined further, there appeared to be an effect of ‘outcome’. Outcome was recorded by NR during the observation to describe any treatment decision or ongoing management of a problem discussed between veterinarian and client.

The type of consultation outcome, the action to be taken regarding treatment going forward, observed and noted by NR appeared to have an effect on the likelihood of a problem being recorded by the vet in the EPR. Outcome effect was considered for 8 categories as defined by NR on the observations record sheets; The data appeared to suggest that an outcome of ‘do nothing’ or to a lesser extent ’continue to manage’ was much less likely to be recorded on the EPR than other outcome category, see Figure 5-6.

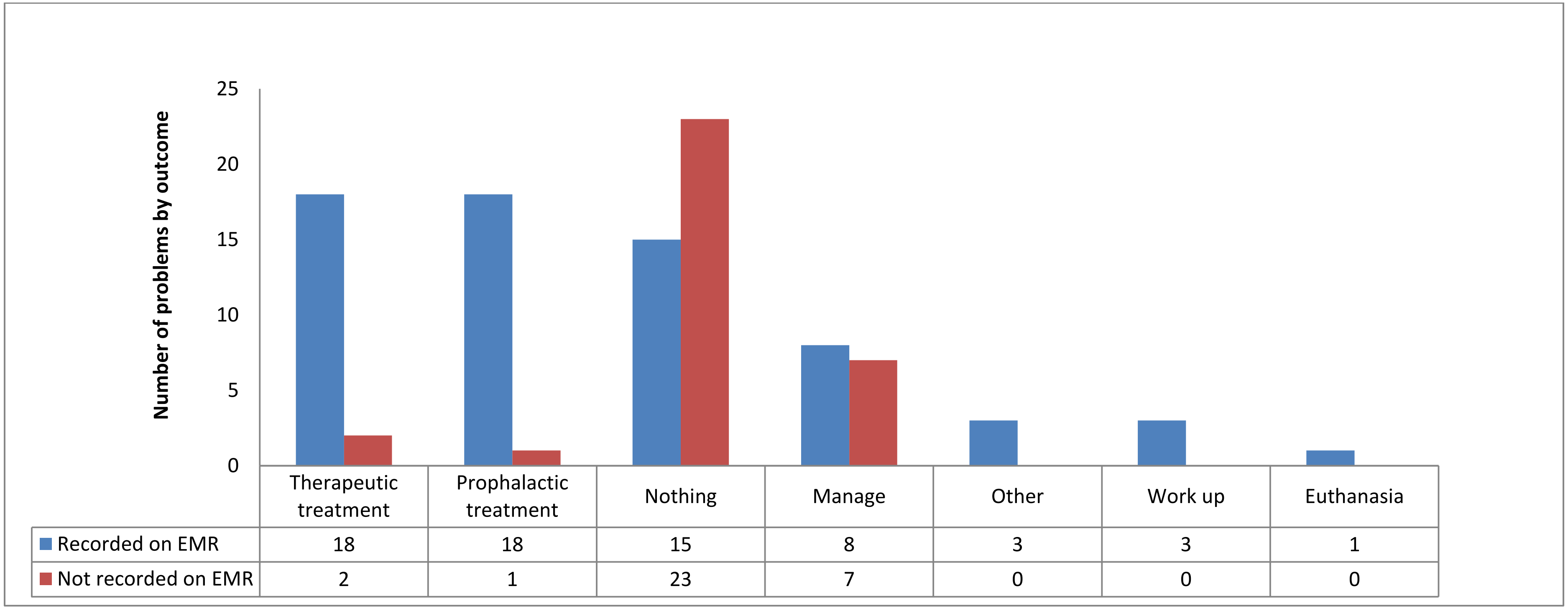


Figure 5-6 Data presenting the analysis of number of records included in the EPR in relation to consultation outcome per problem discussed (n=99).

5.4 Discussion

It is important to understand how accurately an EPR reflects the content of the actual consultation to ensure data collected electronically for research is complete. This study allowed a comparison of two methods of data collection, direct observation and EPR extraction and the measure of agreement between the two was good but not complete.

The sensitivity analysis suggested that where information had been recorded in the EPR it was accurate and comparable to the information collected by observation. However, an important finding of this research was that a number of records were missing from the EPR (35% of all problems discussed). This may be a limitation of data collected by extraction for research and possible highlight areas for future investigation for the research to progress.

The discrepancy appeared to lie in the limited recording of consultations where the outcome was described as 'nothing' or to a lesser extent 'manage'. This highlights differences in use of the EPR by veterinarians compared to the priorities of researchers; where direct observation recorded all discussion points, veterinarians may only use the EPR as an aid memoir for patient care and not a complete record of all events and discussions throughout the day (Bernstein et al., 1993, van Bemmel and Musen, 1997, Shortcliffe and Blois, 2003).

It is unknown why outcomes relating to management or no action were recorded less often than consultations where other outcomes were observed. One possible explanation might be that an outcome of nothing or manage may represent a consultation where no further treatment is required. Another possibility is that where only a question is being asked, possibly unrelated to the current consultation, the veterinarian may not feel the time take to record the discussion is useful or that the information is relevant. The information may already have been recorded elsewhere and therefore be considered repetitious. Alternatively where the details of the problem have not changed from information already captured within the animals record the veterinarian may feel adding repeated information is not good use of time. An example of this latter point may be requests to clip an animals' nails, previous problem discussed but no change in condition apparent, microchip placement check or a recheck of a previous problem with no change in condition.

Results from surveys of the veterinary profession suggest time is probably the most limiting factor to research being conducted in first opinion veterinary practice (Evans et al., 1974). The limitations of time within the consult may play an important role in how much information can be recorded and also affect the type of information that is captured. Where no further progress has occurred in a case or where problems discussed are speculation rather than fact the veterinarian may prioritise deciding this information need not be captured making more time for notes where follow up investigations are likely to be necessary.

Harman et al (2012) examined the use of an EPR in human healthcare by comparing information collected by survey of the healthcare profession with that recorded in the EPR. These researchers found that when the number of problems discussed during the consultation was considered, the patients presenting with 3 or more problems to discuss has significantly different recorded history, and was somewhat limited when compared to those patients with fewer problems to discuss. They also found a relationship between increasing complexity of the problems discussed and recording of consultation outcome.

The findings of the current study did not really identify a similar relationship with the recording of complex verses simple problems, although further research on a larger sample set would be recommended to re-examine this finding. However there was a relationship between the numbers of problems discussed and recorded and this could be an influence of problem number or problem presentation in connection to outcome if the problems a client discusses last are of lower concern and therefore more likely to have an outcome of nothing. It would be interesting to examine this further with a larger dataset.

Discussions with the sentinel veterinarians involved in this study during the feedback sessions have found that time is a factor in how the EPR is completed (*Feedback session minutes, January 2013*). The feedback revealed that one reason why the veterinarians involved in this study may limit details within the EPR is due to the design of the PMS system itself and the time required to complete the input fields in full. It may be that, if the PMS system were designed with greater flexibility, more information may be documented within the EPR during the consultation and also the information that is recorded to be better organised and more complete for use in research.

Although the quality and quantity of the data collected by direct observation far exceeds that of the EPR, direct daily observation for research is impractical as a data collection method long term as it is very labour intensive, time consuming and expensive process. In the human health field alternatives to analysis of the complete electronic patient record, such as clinical coding, have been found to have comparable limitations on the completeness of data (Stein et al., 2000).

The data extraction method, as described previously, is a quick and simple method for gathering clinical data from first opinion veterinary practice. It has very little impact on the veterinarian's time and data can be collected from many veterinarians and consultations simultaneously. Although the results presented here from this sentinel practice suggest only 65% of all information discussed during a week of consultation is captured in the EPR, this research is the first of its kind in veterinary medicine to allow this type of comparison of information captured by direct observation to that captured using electronic data extraction. The findings furthered our understanding of the veterinary consultation and the recording of clinical data by veterinarians in first opinion veterinary practice.

5.5 Conclusion

The disparity between the number of problems discussed during the consultation and those recorded on the EPR was an interesting and important finding, for EPR extraction and analysis for practice-based research. It appears, from the findings presented here, that the outcome of the consultation might play a part in what information is recorded by the veterinarian and therefore warrants further investigation.

Limitations of time within the consult may play a part in how much information can be recorded, this may also affect the type of information captured. Further understanding of limits to recording of information in the EPR could be assessed by running focus groups with veterinarians or interviews. This may be an excellent next step for the future of this research.

6 A review of the British Small Animal Veterinary Association (BSAVA) Congress proceedings abstracts published 2001-2010

6.1 Introduction

In order to guide the subject areas appropriate for veterinary research, it is necessary to identify areas where clinicians perceive there to be the greatest gaps in knowledge (Faunt et al., 2007). These gaps in turn are likely to be related to disease prevalence and incidence within the vet visiting pet population and the conditions presented daily to vets in practice (Holmes, 2007, Vandeweerd et al., 2012c). Information is however still lacking on where to find the information needed to assist with treatment decisions once the gaps in current knowledge are identified.

In the UK the Royal College of Veterinary Surgeons (RCVS) Code of Professional Conduct (Appendix 12) states that all veterinary surgeons “have a responsibility to ensure they continually maintain and develop their knowledge and skills” and it is expected each veterinarian will perform the RCVS recommended minimum number of continuing professional development (CPD) hours; 105 hours over a rolling three year period with an average of 35 hours per year[16]. There are a number of activities which may be considered suitable for CPD to achieve this requirement (Appendix 13) including attending external lectures, presenting a lecture or presentation and conducting research.

The British Small Animal Veterinary Association (BSAVA) has a membership of approximately 8,000 veterinarians and vet nurses and organises and delivers the largest and most popular veterinary congress and trade event in Europe with veterinary attendance reaching 6199 in 2013 (http://www.bsava.com/Membership). Launched in 1958, the first BSAVA Congress was designed as a scientific meeting for invited experts to present papers. In 1961 the 4th meeting was joined by the World Small Animal Veterinary Association (WSAVA) which encouraged attendance by veterinarians from overseas and has been a frequent collaboration since (Singleton, 1999).

BSAVA Congress is a popular meeting for the UK veterinary profession and is a major contributor to veterinary CPD. All material presented relate directly to small animal practice and organisers claim the clinical research abstract

---

[16] www.rcvs.org.uk

sessions represent “The next thing” in small animal research (Ramsey, 2011). The CRA sessions at BSAVA are small satellite meetings with an audience limited to a maximum of 20 or 30 attendees. Anecdotally they generally host researchers wishing to share their most recent research findings and the research can be from a variety of backgrounds such as academia, first opinion referral practice and laboratories. Veterinary stream meetings at BSAVA typically have a large audience of over 100 attendees and cover topics which are considered by members of the congress committee as suitable for the continued professional development (CPD) of veterinary practitioners. The majority of these presentations are given by invited clinical experts. To identify topics relevant for small animal veterinary research and continuing professional development (CPD), it is pertinent to look at subject areas covered during events such as the BSAVA Congress.

This study analysed the veterinary lecture stream (VS) and clinical research abstracts (CRA) published in the British Small Animal Veterinary Association (BSAVA) Congress proceedings over the last 10 years. The main aims of the work were;

1. To investigate the frequency of certain topic areas chosen for VS and CRA presentations over time.
2. To investigate the most popular type of study design used for clinical research presented as CRA over time.

6.2 Materials and Methods

All analysis was performed using published abstracts in the BSAVA Congress proceedings 2001 – 2010.

6.2.1 Abstract content and subject analysis

6.2.1.1 Veterinary Stream Lectures and Clinical Research Abstracts

The individual abstracts for both the VS (n = 1595 abstracts) and CRA (n = 1057 abstracts) sessions were extracted from the proceedings across the 10 years. All abstracts were grouped by session heading as allocated by BSAVA. As many of the session headings were found to be repeats the researcher (*JJD*) identified 111 unique session headings shared by all abstracts (Appendix 14).

The content of each individual abstract (VS & CRA) was then examined by the researcher (JJD) to refine the groups further based on research focus. Clinical research abstracts and veterinary stream abstracts were considered separately. The session headings were refined from the 111 unique session headings into a smaller group of 23 final headings e.g. Abdominal imaging, Imaging, Thoracic imaging, Musculoskeletal imaging are all unique session headings and were all placed into one group of abstracts with the final heading Diagnostic imaging. The final 23 session headings are shown below in Table 6-1.

Table 6-1 List of final group allocation for BSAVA VS & CRA abstracts 2000- 2010

Number	Details
1	Anaesthesia & Critical Care
2	Cardiology, Vascular & Respiratory
3	Clinical Pathology
4	Dentistry
5	Dermatology
6	Diagnostic Imaging
7	Endocrinology
8	Epidemiology, Survey & EBVM
9	Exotics
10	Genetics
11	Hepatology, Gastroenterology & Nutrition
12	Infection & Immunity
13	Nephrology & Urology
14	Neurology
15	Oncology
16	Ophthalmology
17	Orthopedics
18	Rabbits, Birds & reptiles
19	Reproduction
20	Soft Tissue Surgery
21	Therapeutics
22	Welfare, behavior & Ethics
23	Other

Each abstract was then re-allocated to one of 23 final group headings by JJD. Four meetings were held in total to confirm allocation of abstracts and three members of the Centre for Evidence-based veterinary medicine (CEVM), JJD, MB and RD, were consistently present at all meetings. The allocation of abstracts was confirmed or disputed by the CEVM group discussion and a consensus reached for final allocation of abstracts. Where an abstract could not be grouped into a final heading it was placed into a group for miscellaneous abstracts termed 'other'.

Where a session heading was entitled ‘Top tips for…..’ or ‘…..made easy’ and therefore could not be allocated to any single specific heading based on the abstract category the abstracts were separated and allocated to a relevant group by *JJD* and the team based on abstract content.

Where a session heading was species specific, such as the ‘cat group session’ or ‘feline forum’, the content of the abstract (e.g. body system/veterinary discipline) rather than the species was used for allocation. However where an exotic or wild animal was the focus of the abstract, the abstract was allocated to the ‘Exotics’ group as many exotics abstracts covered whole animal information and were not body system specific.

The final grouping of the exotics was found to be extensive so were further separated into two groups: **‘Rabbits birds and reptiles’** and **‘Exotics’.** The Exotics’ category included any exotic species which did not fit into the rabbits, birds or reptiles category. Is there a reason for nothing else on this page Jules?

6.2.2 Review of Clinical Research Abstracts

6.2.2.1 Review criteria

A review of a subset of CRA abstracts was performed to assess the type of study most frequently published for the CRA stream and to gauge the type of research presented at the Congress over the last decade. The diagram in

Figure 6-1 demonstrates typically how research studies are categorised in relation to the type of information, or strength of evidence, they provide.

Abstracts for each year (2001 – 2010) were chosen for review. Using a Microsoft Excel random number generator function, a 10% random sample from the total number of abstracts per year of the congress was selected. Two studies were performed; a pilot study of a 10% sample from three years of abstracts (2001 – 2003) to initially test the review method, and a research study using a 10% sample from ten years of abstracts (2001 – 2010) to perform the actual review.

Systematic Review
Meta-analysis

Experimental
Randomised Controlled Trial (RCT)
Controlled or Randomised Clinical Trial

Observational
Cohort
Case Control
Cross Sectional study or Survey

Descriptive
Case Series
Single case report

In-vitro
Clinical trial
Diagnostic methods
Unclassifiable/ Discovery research

Figure 6-1 List of study type presented in decreasing order for research evidence according to (Kastelic, 2006, Lean et al., 2009).

A method to identify the study design described by each CRA was designed as follows,

i. The descriptions for each CRA abstract were considered and a definition for each design was agreed by members of the CEVM (Appendix 15).
ii. To test the chosen definitions were suitable for use, a pilot study was designed. A consensus of three researchers from the CEVM (JJD, MD & JS) was reached on the final study type during CRA review applying the study definitions.
iii. Once validated the process was repeated as a research study for a 10% sample of all the CRA from 2000-2010. The first three years 2001-2003 were included in the review again however the randomisation was repeated to create a new selection of abstracts for review.
iv. If it was unclear which study design had been used, the review was repeated and study design decided by JJD and an independent consensus of 2 different researchers from the CEVM (MB, RD). The review was repeated until all three researchers agreed on the study design to allocate and all CRA were allocated.
v. After 4 review sessions; where the study design could not be allocated by the three reviewers into any of the groupings the abstract was allocated to an unclassified grouping.

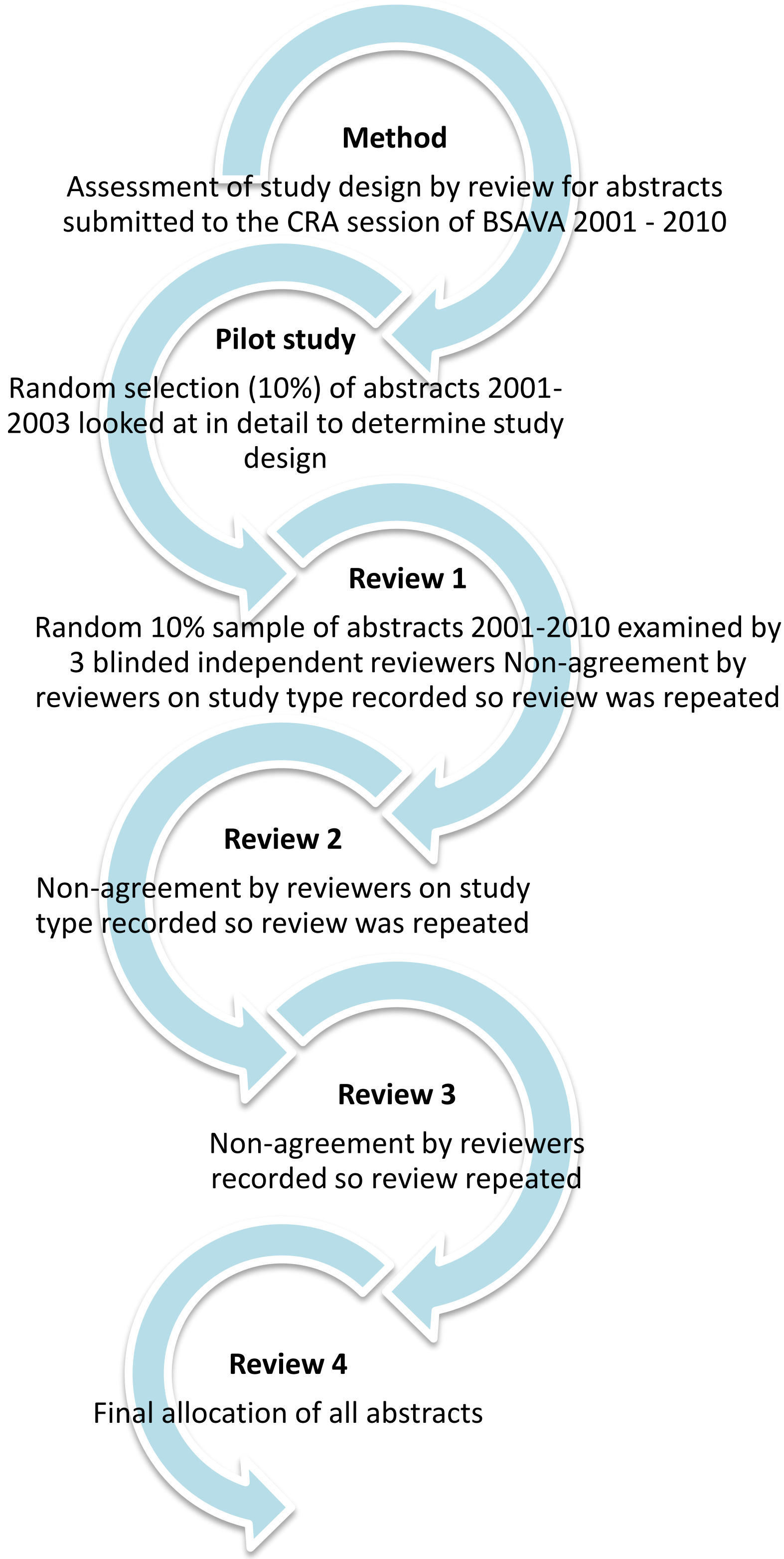


Figure 6-2 Flow diagram describing method used for abstract review and study type allocation

The study design definitions were mostly taken from Kastelic (2006) except for some slight variations where the study was not one of the more common types of study design (highlighted below). The full list of study types considered during the review was as follows;

1. Systematic reviews and Meta-analysis
2. Randomized controlled trial
3. Controlled clinical trial
4. Randomised Clinical Trial
5. Cross-over Clinical trial
6. Observational studies - Cohort studies
7. Observational studies - Case-control study
8. Observational studies - Cross-sectional survey
9. Comparative Research
10. Clinical research trial
11. Case series
12. Single Case reports
13. Expert opinion
14. In-Vitro
15. Diagnostic Methods
16. Unclassified, discovery research.

Each reviewer allocated each study to a particular category blind to the decisions of the other reviewers. A decision was made on study type where at least two researchers agreed. Once all the abstracts had been allocated to a study design an analysis was performed to examine,

i. The different study designs published within the BSAVA CRA conference proceedings.
ii. An analysis of study design frequency over time to investigate any change in study type year to year over the 10 years.

6.3 Results

6.3.1 Abstract content and subject analysis

There were a total of 111 unique session headings describing both VS & CRA sessions, over the 10 year period (Table 6.2). In some cases, more than one session per year had the same session heading to provide opportunity for a larger number of presentations; for example in 2001 there were 15 Orthopaedic abstracts within the VS and therefore 2 sessions published in the proceedings for that year.

6.3.1.1 Veterinary Stream Lectures

The BSAVA Congress proceedings from 2001 – 2010 contained a total of 272 VS lecture sessions with 1595 abstracts (Table 6.2). There was a mean of 27 VS sessions per year (range min 23 - max 32). For each session there was a mean of 6 abstracts published, however this varied year to year (VS mean 6, min 1 – max 9, Table 6.2).

The most frequent session heading for abstracts within the VS over the last 10 years was initially found to be the exotics group (n = 146, min 7 max 24 per year). Following the further separation of the exotics session into two smaller groups i) Rabbit, bird & reptile (n=99), ii) Exotics (n=47), the frequencies were adjusted. Cardiology, Vascular & Respiration (n = 144, min 5 max 20 per year) was then the most presented session heading within the veterinary stream followed by Orthopaedic (n = 127, min 8 max 15 per year) and Hepatology Gastroenterology & Nutrition (n = 116, min 6 max 18 per year, Table 6.2).

6.3.1.2 Clinical Research Abstracts

The BSAVA published a total of 169 CRA sessions over a 10 year period 2001 – 2010 with a total of 1057 Clinical Research abstracts in their Congress proceedings. A mean of 17 CRA sessions were published per year (min 15 - max 20) in the proceedings. Within each session there was an average of six abstracts published, again this varied year to year (CRA mean 6, range 1-8 per session, Table 6-3).

Orthopaedic abstracts were the most popular publication in the CRA (n = 156, min 11 – max 23), followed by Hepatology Gastroenterology & Nutrition (n = 116, min 6 – max 26) and Infection & Immunity (n=103, min 1 – max 14).

Table 6-2 Number of abstracts published per session in BSAVA Congress proceedings Veterinary stream (VS) 2001 – 2010

	Veterinary Lecture Stream Abstracts	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	Total per subject over 10 years
1	Cardiology, Vascular & Respiratory	20	16	13	18	20	6	5	17	14	15	144
2	Orthopaedics	15	13	14	15	15	12	12	9	8	14	127
3	Hepatology, Gastroenterology & Nutrition	8	18	9	14	14	6	13	10	13	11	116
4	Clinical Pathology	9	14	6	7	13	7	16	7	12	12	103
5	Anaesthesia & Critical care	11	6	15	2	6	15	20	6	10	8	99
6	Rabbits, Birds & reptiles	12	14	14	11	12	8	14	4	5	5	**99**
7	Dermatology	8	10	14	9	15	9	4	7	10	6	**92**
8	Welfare, Behaviour and Ethics	0	6	12	12	12	13	8	8	4	4	**79**
9	Oncology	0	7	6	6	8	8	15	10	7	8	**75**
10	Diagnostic Imaging	7	8	4	0	1	16	0	13	7	17	**73**
11	Soft tissue surgery	4	1	12	5	12	12	6	6	9	6	**73**
12	Endocrinology	6	6	8	6	9	8	1	7	12	6	**69**
13	Nephrology & urology	10	7	7	10	4	8	9	2	2	8	**67**
14	Ophthalmology	7	6	8	4	7	6	6	8	6	8	**66**
15	Neurology	9	7	5	17	7	6	0		8	5	**64**
16	Infection & immunity	8	9	0	14	2	10	7	8	0	2	**60**
17	Exotics	0	10	4	3	4	8	0	3	5	10	**47**
18	Dentistry	0	0	7	4	6	6	7	6	1	4	**41**
19	Other	5	3	0	7	0	4	6	3	6	4	**38**
20	Reproduction	6	0	0	0	6	0	0	8	0	1	**21**
21	Therapeutics	3	0	6	0	2	0	3		4	1	**19**
22	Epidemiology, Survey & EBVM	4	0	0	0	6	0	0	6	1	0	**17**
23	Genetics	0	0	0	0	0	0	4		1	1	**6**
												1595
	Presented Abstracts/ Year	**152**	**161**	**164**	**164**	**181**	**168**	**156**	**148**	**145**	**156**	**1595**
	Sessions/Groups in total	**23**	**27**	**28**	**30**	**32**	**31**	**27**	**24**	**25**	**25**	**272**
	Abstracts (including those withdrawn)	**152**	**161**	**164**	**164**	**181**	**168**	**156**	**148**	**145**	**156**	**1595**
	Number of abstracts withdrawn	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**

Table 6-3 Number of abstracts published per session in BSAVA Congress proceedings Clinical research abstracts (CRA) 2001 – 2010

	Clinical Research Abstracts	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	Total per subject over 10 years
1	*Orthopaedics*	19	14	17	20	23	12	18	11	11	11	156
2	*Hepatology, Gastroenterology & Nutrition*	10	12	7	12	10	6	7	8	26	18	116
3	*Infection & immunity*	1	12	7	14	12	12	12	11	8	14	103
4	*Oncology*	7	12	7	6	9	18	10	7	8	6	90
5	*Ophthalmology*	7	10	10	7	10	6	6	5	6	6	73
6	Neurology	8	7	5	9	6	6	10	11	4	6	**72**
7	Endocrinology	5	5	10	7	9	7	4	8	2	3	**60**
8	Soft tissue surgery		5	14		5	6	11	7	6	6	**60**
9	Cardiology, Vascular & Respiratory	6	7	6	6	6	3	2	6	1	11	**54**
10	Diagnostic Imaging		4	7	4	8	7	8	6	6		**50**
11	Anaesthesia & Critical care		3	7	7	8	4	5	1	3	4	**42**
12	Dermatology	11	3	3	6	4	4		3	2	4	**40**
13	Welfare, Behaviour and Ethics	6	7	5	3	6	4	1		1	6	**39**
14	Clinical Pathology	3		2	4	4	5	3	5	3	6	**35**
15	Nephrology & urology	7		3			2	1	4	2	6	**25**
16	Therapeutics		3	4					5			**12**
17	Rabbits, Birds & reptiles	7					3					**10**
18	Exotics							3	2	2	1	**8**
19	Genetics						1	4				**5**
20	Other	1		3								**4**
21	Epidemiology, Survey & EBVM									3		**3**
22	Dentistry											**0**
23	Reproduction											**0**
												1057
	Presented Abstracts/ Year	**98**	**104**	**117**	**105**	**120**	**106**	**105**	**100**	**94**	**108**	**1057**
	Sessions/Groups	**15**	**15**	**18**	**15**	**15**	**20**	**18**	**18**	**18**	**17**	**169**
	Abstracts (including those withdrawn)	**102**	**105**	**118**	**105**	**121**	**107**	**105**	**100**	**94**	**108**	**1065**
	Number of abstracts withdrawn	**4**	**1**	**1**		**1**	**1**					**8**

6.3.1.3 Overview

The five most frequent subject areas chosen for publication in VS and CRA were the same: Orthopaedics, Hepatology Gastroenterology & Nutrition, Oncology, Cardiology Vascular & Respiration and Diagnostic Imaging. Some topics, e.g. Dentistry, featured in the VS lectures for 8 consecutive years (2003–2010), but never in the CRA. Similarly Reproduction featured occasionally within the VS but never in the CRA (Figure 6-3).

There were annual variations in the frequency with which abstracts were presented for both VS and CRA sessions over the 10 year period. The ten most popular session headings across the 10 years for VS sessions are shown in Figure 6-4. The data for VS was compared to the CRA sessions for the same subject areas as total number of abstracts (CRA & VS) published as a proportion (%) of all ten session areas combined (100%). It clear that there was greater variation within the number of abstracts presented per CRA session per year when compared to the VS lectures which remained relatively constant across the 10 years.

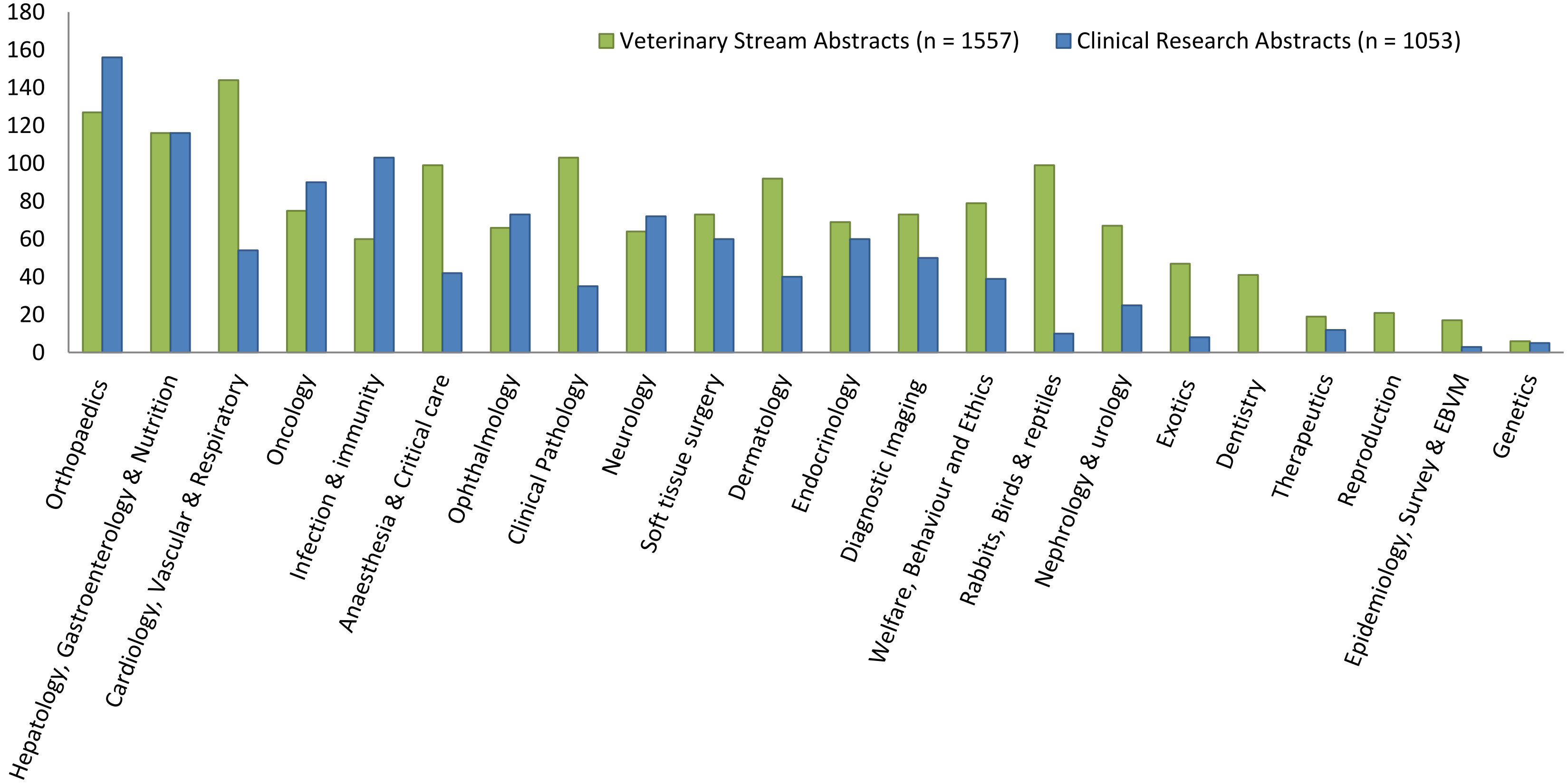


Figure 6-3 Total number of abstracts (excluding other category) published for VS lectures and CRA from 2001-2010 in BSAVA Congress proceedings.

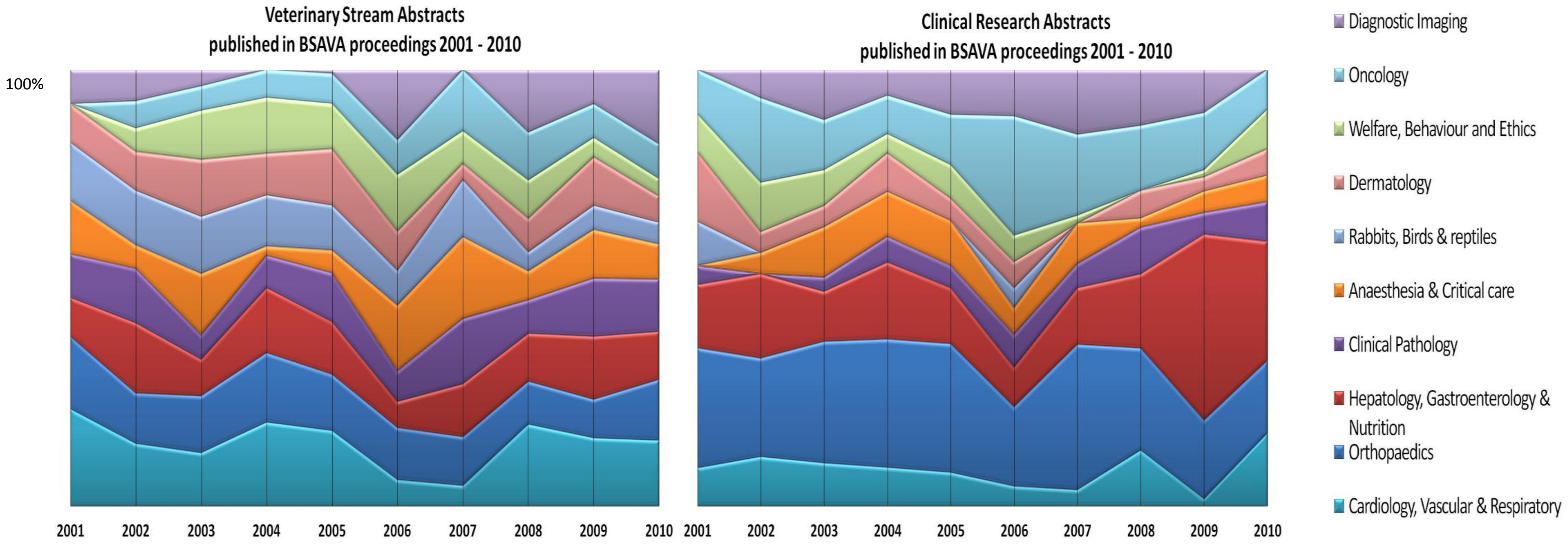


Figure 6-4 Comparison of the ten most common VS session headings compared to the same CRA headings.

6.3.2 Review of Clinical Research Abstracts

6.3.2.1 Pilot Study

The pilot was successful and all abstracts could be allocated to a single study design using the parameters set for the review.

6.3.2.2 Research Study

6.3.2.2.1 Study Design

The study types in the abstracts published in the BSAVA congress proceedings were not easily identified during the review process. After 4 review sessions the abstracts were finally allocated to one of 10 study designs out of a possible 16 (Figure 6-5). The graph in

Figure 6-6 represents the final allocation of 107 abstracts by study design. Case series is by far the most frequently published study type n = 33 (31%), followed by studies designed for in-vitro research (15%) and diagnostic methods (15%). There were a small number of cross-sectional studies (9%), Randomised Controlled Trials (6%) and clinical trials using a control group (5%). However there were no cohort studies, cross-over clinical trials or randomised clinical trials identified.

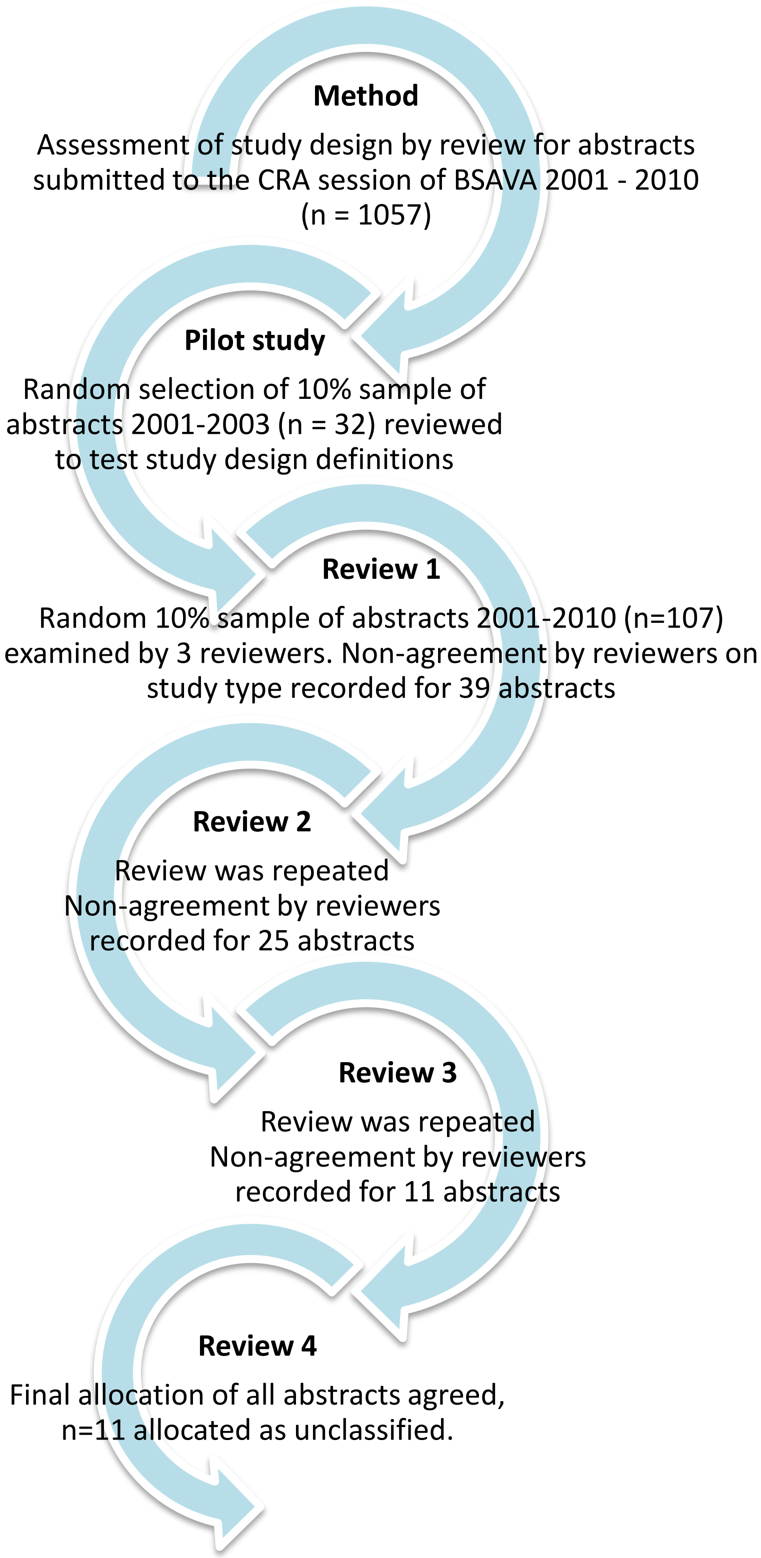


Figure 6-5 Flow diagram describing results for abstract review

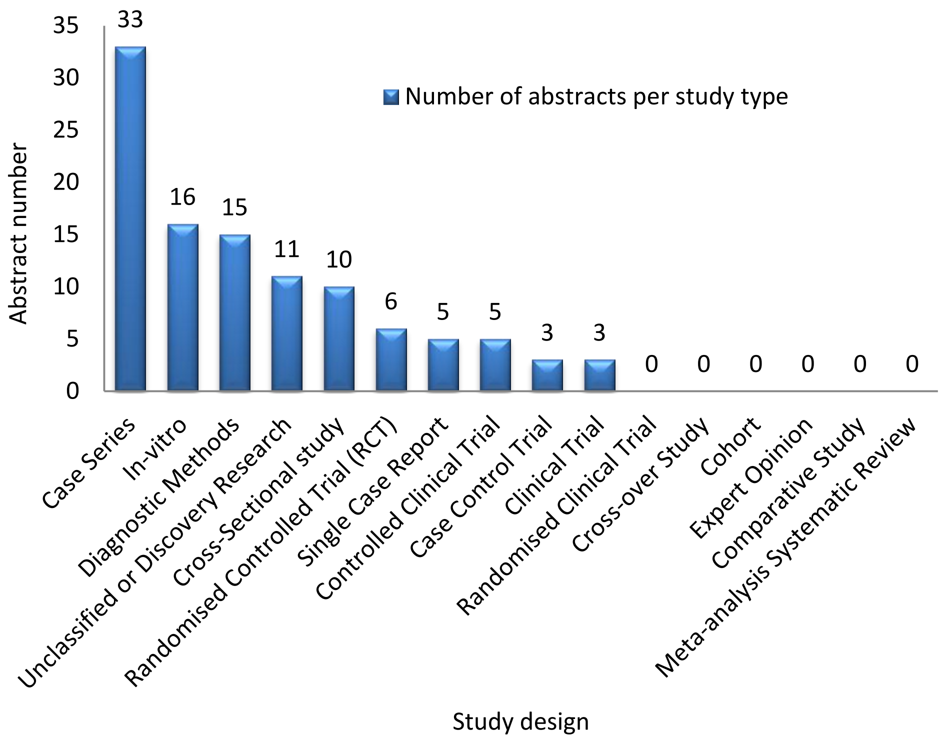


Figure 6-6 Study type frequency of a sample of the BSAVA CRA abstracts published in Congress proceedings over a 10 year period (n= 107).

6.3.2.2.2 Analysis of study type, year to year, over 10 years.

The results are presented in full in Table 6.4. The results suggest the number of RCTs, although still very low, almost doubled in 2010, the most recent year of publication of this review, when compared to the other 9 years. In addition the publication of Controlled Clinical Trials, Case Controls and Cross Sectional research all increased in frequency in the final three years of the review when compared to previous years. The number of case series, whilst the most frequently study type throughout all years, was present in a much lower frequency in the last year of the review (2010).

Table 6-4 Study design data for each year of BSAVA Congress

Year	**2001**		**2002**		**2003**		**2004**		**2005**		**2006**		**2007**		**2008**		**2009**		**2010**	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Case Series	4	40.0	3	30.0	5	42.0	3	27.2	4	33.3	4	36.4	3	27.2	3	30.0	3	33.3	1	9.1
In-Vitro	1	10.0	3	30.0	1	8.3	1	9.1	2	16.7	4	36.4	2	18.3	0	0	0	0	2	18.3
Diagnostic Methods	3	30.0	1	10.0	1	8.3	3	27.2	3	25.0	1	9.1	0	0	1	10.0	1	11.1	1	9.1
Unclassified	0	0	0	0	1	8.3	2	18.3	2	16.7	0	0	2	18.3	2	20.0	1	11.1	1	9.1
Cross-Sectional study	1	10.0	0	0	1	8.3	1	9.1	0	0	1	9.1	2	18.3	1	10.0	2	22.2	1	9.1
RCT	0	0	0	0	1	8.3	1	9.1	1	8.3	0	0	1	9.1	0	0	0	0	2	18.3
Single case report	1	10.0	1	10.0	0	0	0	0	0	0	1	9.1	0	0	1	10.0	0	0	1	9.1
Controlled Clinical trial	0	0	1	10.0	1	8.3	0	0	0	0	0	0	0	0	1	10.0	1	11.1	1	9.1
Case Control trial	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	10.0	1	11.1	1	9.1
Clinical Trial	0	0	1	10.0	1	8.3	0	0	0	0	0	0	1	9.1	0	0	0	0	0	0
Count	**10**		**10**		**12**		**11**		**12**		**11**		**11**		**10**		**9**		**11**	

6.4 Discussion

An analysis of the VS and CRA abstracts published in the proceedings of the BSAVA Congress over the last decade allowed an insight into the type of CPD and research presented. It is assumed from the data that the priorities of the congress 'veterinary stream' focus on much the same subjects from year to year with very little variation. The clinical research abstracts show a much greater variation in focus from year to year but in the abstracts that were assessed, this 'new' research appears to be somewhat unrelated to what is delivered in the VS.

Experimental trials providing evidence considered to be most reliable to clinical decision making, such as randomised controlled trials and controlled clinical trials, were rarely identified within the CRA abstracts of BSAVA. In comparison the large number of case series and single case reports, while still of value, present research typically more useful for hypothesis development. BSAVA is a practice-based congress and as such is more likely to attract case reports or case series which are possibly easier to prepare and present for a busy veterinary practitioner. Large scale practice-based research is relatively new. It is possible; as more practice-based research is conducted this trend may change in time. However the high proportion of case series coupled with limited numbers of experimental (RCT, cross-over studies) or observational

(Cohort, cross-sectional or case/control studies) studies may have implications for the evidence base, or lack of, available to practitioners.

If the VS lectures, and therefore the CPD they offer, are ‘evidence-based’ from large scale practice-based research then over time the VS lecture content may be closer aligned to the CRA content. In addition subjects popular for CPD should drive the areas for research and eventually connect the two together with a combined aim of translating research back into practice. The findings of this study could suggest this has not happened over the last decade, although the analysis from year to year shows a promising change in this trend over the last few years with the numbers of RCTs, case-control studies and controlled clinical research chosen for presentation all increasing from 2008.

A possible limitation of this research is that research findings may be presented elsewhere. Specific clinical speciality research meetings e.g. British Veterinary Dental Association[17], could be a preferred forum and therefore are not being fed into the VS at BSAVA. It is also possible that a preferred choice for CPD is subject specific conferences, or dedicated CPD courses, rather than BSAVA Congress.

---

[17] http://www.bvda.co.uk

A second limitation is that original research papers were not used for the classification of study design and only the abstracts published in the proceedings. There are very few examples of published research considering the overall scope of research presented at meetings and CPD events, although see Boden et al., (2010). The completeness of the results published in abstracts for outcomes research has been examined by Snedeker et al., (2010) who found there may be a disparity between the information provided by an abstract and subsequent peer reviewed publication of the corresponding research paper. Therefore this should be a consideration when interpreting the results presented here.

Finally, this study only selected a 10% sample of abstracts for convenience to classify study types which may not be representative of the total body of abstracts. Additionally the small sample led to a subsequent low sample of abstracts in each group reducing the power of the analysis. The purpose of this exercise was to generate hypotheses rather than prove a hypothesis; therefore, whilst a power calculation may have been useful to quantify the exact limitations of the small sample size, these calculations were not carried out as it was not thought to be necessary to identify the required number of abstracts in order to draw definitive conclusions.

It is possible that if full papers were interpreted rather than abstracts, further detail on the research methods used could have revealed different study types to those identified More detail such as that found in full papers would have provided fuller understanding of which research methods had been applied and why.

It is likely that topics for the VS abstracts are determined in advance of the congress, before the CRAs have been selected which could affect the results found here. It is also unknown how many first opinion practitioners attend or submit to the CRA stream in comparison to the VS streams which could have a bearing on the relationship between the two types of abstracts.

Finally the system of abstract selection and allocation for Congress is conducted by the BSAVA Congress committee[18] (*pers.comms* Helen Field March 2011). It is unclear if the selection procedure is the same from year to year, as the Congress committee changes regularly. It is also unknown if what is selected for presentation at Congress is what the veterinary profession would select, and how this relates to what is submitted and not selected, or what is being submitted elsewhere. Further research into the attendance by vets of BSAVA Congress for CPD would contribute valuable additional information.

---

[18] http://www.bsava.com/BSAVACongress/MeettheTeam

6.5 Conclusion

This study allowed comparison between selections of the two types of presentation delivered at BSAVA over the last decade and may be a starting point to help identify areas where greater overlap is needed to connect veterinary research with veterinary practice. Further work identifying the significance of this relationship to the knowledge veterinarians' gain from attending the various streams at the conference is required to draw definitive conclusions.

Chapter Seven

7 Discussion

7.1 Overview

An evidence-based approach to veterinary medicine should include an understanding of how to appraise scientific evidence, how to identify answers to clinical questions in a systematic way and how to apply the new found information accordingly to support a clinical decision (Greenhalgh, 2010).

The review of BSAVA abstracts was performed to investigate the frequency with which certain topic areas are chosen for Veterinary Stream lectures (VS) and Clinical Research Abstract (CRA) presentations over time and to investigate the most popular type of study design used for clinical research presented as CRA. A critique of the research presented is only one side of the story and what is perhaps more important is how the information may be received, evaluated and applied by the veterinary profession (Slater, 2010).

Veterinary practice-based research is a challenging but valuable task (Cockcroft and Holmes, 2003, Faunt et al., 2007, Slater, 2010) and may be able to provide much needed information to support clinical decision making. The undergraduate teaching of evidence-based methods and critical appraisal of research to veterinary students is vital in providing newly qualified practitioners with the skills and knowledge required to critically appraise and apply clinical research findings. Although opportunities for research in practice may be somewhat limited by the many demands placed on a veterinarian's time (Evans et al., 1974). It is important that the need is still recognised by veterinarians in practice and met by the undergraduate syllabus.

The approach presented here, in the experimental chapters, will hopefully overcome some of the barriers such as time or opportunity to practice-based research. The application of information extraction methods to access and analyse electronic patient records (EPRs) from first opinion veterinary practice will hopefully provide more opportunities for practice-based research and clinical epidemiology and support the move towards more evidence-based veterinary practice. The work presented demonstrates a viable and practical method for data extraction and the information collected has the potential to provide answers to questions of disease prevalence, incidence and treatment outcomes for the vet-visiting pet population in the United Kingdom.

Designing efficient ways of accessing the data collected by vets on a daily basis, specifically the free text, and extracting this information from the different PMS systems would enable vast amounts of clinical data to be collated and analysed. Until standard methods of recording data between and across practices are in place; including veterinary terminology standards (Case et al., 2000), clinical coding systems (Lund et al., 1998, Hahn et al., 2003, VeNom coding group) or recording standards for clinical information in electronic patient records (Bernstein et al., 1993, Ferranti et al., 2006, Häyrinen et al., 2008), this extraction and content analysis method will be an invaluable, and currently the only, tool for practice-based epidemiology research and evidence-based veterinary medicine.

7.2 Summary of Findings

7.2.1 Veterinary Informatics and Practice-based Research

This section of the thesis outlined the establishment and validation of a method for the extraction of patient records from a computer based veterinary practice management software (PMS) system for use in practice-based-research.

To extract all target data recorded in a typical electronic patient record (EPR), an Extensible Mark-up Language (XML) Schema (Katehakis et al., 2001) and associated extraction method was designed, as presented in **Chapter 2.** Twenty one data fields were identified for extraction and included all animal demographic fields, consultation information and any treatments given. The Vet-One PMS system was used for a pilot study to validate the use of an XML schema for EPR extraction. The XML schema was then integrated into a working sentinel practice Vet-One PMS system. A database was also designed for the safe storage of extracted patient data.

The XML schema design and EPR extraction method extracted all target EPRs within the pilot CEVM PMS validated by a comprehensive manual check against 326 paper EPRs. The XML schema integrated into the sentinel PMS successfully and validation found the method to have a 100% level of accuracy when compared to a 10% convenience sample (n = 253) of paper records. The data warehouse managed the extracted records securely and all queries within the database tables separated, and filed the data efficiently. The success of the methodology and the high level of precision for the extraction system provide great encouragement for the future of this work and suggest utilising the electronic patient record is a viable method for applying clinical epidemiology in support of evidence-based veterinary medicine and practice-based research.

A method of content analysis and a system of text mining of clinical information extracted from a veterinary EPR was discussed in **Chapter 3** using a keyword in context (KWIC) method and also a keyword retrieval method both facilitated by the use of WordStat software. The keyword in context function was applied to the CEVM pilot EPR to text mine the clinical information. An EPR extracted from the collaborating sentinel practice was also examined using the WordStat KWIC and a keyword retrieval (KR) facility to investigate the suitability of the software for use on real patient information. Finally the text mining method was validated by comparison to a manual examination of extracted clinical information.

The results of the CEVM pilot EPR study showed a very high level of accuracy for content analysis using KWIC. Using the sentinel practice EPR, the KR function out-performed the KWIC search and was found to have a higher level of positive predictive value (PPV=100%) than the KWIC search (PPV=94%). The results found the use of WordStat software and the two methods of text mining used for this research, KWIC and Keyword Retrieval, offered an excellent level of success. Validation using data collected from the sentinel EPR produced a higher level of success for the KR method. The results of the study suggest the method would be a useful resource for the analysis of information extracted from PMS system EPRs.

In **Chapter 4** a new and unique text mining methodology was introduced, designing a dictionary of disease specific terms to analyse electronic patient records extracted from a first opinion sentinel veterinary practice. The dictionary was created to identify patients which share a common clinical condition; feline lower urinary tract disease (FLUTD). The information contained within a veterinary practice management software system electronic patient record was used to develop the methodology and to test the efficacy of a computer assisted free text analysis program (WordStat) to mine electronic patient records using a dictionary of clinical terms.

The dictionary search of the EPR revealed many of the terms selected to create the dictionary were not used by the veterinarian to record cats with FLUTD and only a small proportion (4%) of the selected dictionary terms were actually present in the EPR. However regardless of this lack of information Wordstat provided an excellent level of precision for search and retrieval of FLUTD patient records from within a sentinel practice EPR. The program also provided a feedback function which allowed the dictionaries to be further refined to create an optimal search dictionary. The method described here offers a quick and efficient means to identify a high proportion of positive cases in a relatively short timeframe and provides and excellent solution to the current difficulties for practice-based research.

The introduction of computers to veterinary practice is thought to offer an improvement in the organisation and efficiency of patient care (Foley et al., 2005) and could be a useful tool for practice-based research. However, the conversion from paper to a computerised record system has primarily been to support the clinical care of patients, and not necessarily research. Therefore it is not known how much clinical information is actually recorded in the EPR and how the information that is captured relates to the details discussed and treatment decisions made between the veterinarian and the client during the consultation. The aim of the final experimental chapter, **Chapter 5**, was to compare the information captured and extracted from a practice management software system, during one week of consultations with information collected by direct observation of the veterinary consultation. Analyses included comparison of information captured for each consultation by both methods of data collection (observation and EPR extraction). The method for comparison of data collection was facilitated by veterinary review. All consultations were performed by 2 veterinarians. Observations made in real time included visits by 36 animals with 99 problems recorded. The same animal data recorded in the EPR, when compared to direct observation, was found to represent only 65% of all problems discussed.

The disparity between the number of problems discussed during the consultation and those recorded on the EPR is a concern. However the findings from direct observation suggest only consults where 'manage' or 'nothing' is the outcome are absent from the ERP. Limitations of time within the consult may play a part in how much information can be recorded and may also affect the type of information captured.

7.2.2 A Decade of BSAVA – Research and CPD

An understanding of the research priorities of the veterinary profession was required for the development of the methods presented in the experimental chapters of this thesis. This was attempted and the results were presented in **Chapter 6**. An analysis was performed of the veterinary stream (VS) and clinical research abstracts (CRA) published in the BSAVA Congress proceedings over the last 10 years, to provide some insight into trends in subject areas covered and therefore research priorities.

The priorities of the congress VS were found to focus on much the same subjects. However the clinical research abstracts show a much greater variation. The research presented in the CRA was predominantly case reports and case series, understandably as the easiest form of data to collect and report for individual veterinarians in practice is currently on a case by case basis. In addition BSAVA Congress is primarily a practice-based meeting attracting case-reports and case series and less likely to attract RCTs or experimental studies than more specialised or academic meetings. However this could suggest BSAVA Congress may not be meeting the needs of the veterinary professional for professional development.

The findings do therefore highlight a possible cause for concern regarding the research and practical continued professional development (CPD) from presentations delivered at BSAVA Congress. The high proportion of case series coupled with limited numbers of experimental (RCT, cross-over studies) or observational (Cohort, cross-sectional or case/control studies) studies may have implications for the evidence base, or lack of evidence-base, presently available to clinicians wishing to introduce evidence-based ways of working into their practice (Slater and Boothe, 1995, Cockcroft and Holmes, 2003, Slater, 2010). It is also possible of course that a preferred forum for CPD is subject specific conferences or dedicated CPD courses, rather than BSAVA Congress. Therefore this cannot be a definitive conclusion and warrants further investigation, perhaps at other CPD events or subject specific veterinary conferences.

7.3 Limitations

The method presented in the previous chapters 2-6 was rigorously tested and found to be valid and precise. There is of course limitations with medical informatics, in particular the validity or completeness of information held within the EPR for both human and veterinary clinical informatics. This research was no exception and the differences identified between the information collected by direct observation and that recorded and extracted from the EPR require further investigation. Recent research (Harman et al., 2012) has suggested the recording of clinical information in first opinion (medical) practice may actually be less effective with the introduction of computerisation and PMS systems. It is suggested the combination of coding and data entry combined with limited consultation time restricts the amount of information which can be captured and, as a result, information may be lost. Data quality has also been found to vary. Common causes for information error in the EPR, taken from Bernstein (1993), are (i) human error, physicians entering information incorrectly or not at all, (ii) laboratory or practice personnel errors during the input or transcription of results or records, (iii) misinformation from patients and (iv) concerns over loss of confidentiality, which may limit the completeness of the information captured. All of these causes are difficult to control and may in fact be confounded by factors such as excessive and unpredictable workload or restrictions of time owing to the common practice of a 10 minute consultation (Evans et al., 1974, Bernstein et al., 1993, Everitt, 2011).

It is possible all the factors presented above for errors in the medical EPR may well be the same for the veterinary EPR, although this was not examined as part of this study. However feedback from the sentinel veterinarians involved in this study commented that the time taken to complete the fields in the PMS possibly caused them to limit the information recorded.

A recent study of the medical EPR in the US (Harman et al., 2012) found that information was missed from primary care consultation records when more than two chronic clinical conditions had been discussed during the consultation. It has been suggested the relationship between GP and patient may be disrupted due to the GP-computer interaction and attention given more to recording of information. The data that appeared to be missed in this study was found to be observational cues for behavioural or mental health disorders possibly as a result of practitioners required to record multiple signs into separate note fields within an EPR.

The result of the study presented here suggest a very different finding, where the information recorded less often was not related to the number of problems presented but more likely related to the treatment outcome of the consultation. This finding is interesting and the recording of information in the EPR by veterinarians will be a focus for future research. However sample size is a limitation of this research as only one week of data was collected by observation and therefore the dataset was limited. It would be valuable to repeat the work with a larger set of observation data for comparison to that extracted, particularly re-examining number of problems presented at each consultation and the order of presentation by the client.

The analysis, and to some extent extraction, of medical information from the EPR requires a comprehensive knowledge of human linguistics, in particular medical language and the semantics and synonyms of medical terms. In addition, the lack of standardisation in veterinary clinical terminology as previously discussed further complicates the task. A limitation of medical informatics, for data collection, is that it requires specialised knowledge and skills in both linguistics and clinical epidemiology to not only access the information but also understand the findings. Therefore the method is not easily applied to the practical setting without a dedicated research effort.

Finally, limited access to high quality research is a major hurdle to evidence-based veterinary medicine (Slater, 2010). In chapter 6 the disparity between the types of research presented as CRAs and the strength of evidence they provided was highlighted. However, as discussed it is possible this is not a representative sample of research currently undertaken in small animal practice.

7.4 Discussion

Tools which facilitate evidence-based veterinary medicine (EVM), with minimal disruption to the work of the veterinarian, are crucial to EVM implementation and success. The practice of EVM requires a continued commitment and a combination of clinical expertise, a consideration of the patient needs and circumstances and an understanding of the best available and applicable evidence to support clinical decision making (Schmidt, 2007, Cockcroft and Holmes, 2003).

The research designed and presented here can contribute to clinical veterinary epidemiology and EVM by providing high quality evidence collected from patient centred research and well designed and statistically robust population based research findings. By collecting the data remotely and synthesising this information into a user friendly format it is possible to support veterinarians to conduct practice-based research whilst reducing the need for specialist skills in clinical epidemiology. However, for this to support the veterinary practitioner its value at least needs to be understood by the practicing veterinarian (Slater, 2010).

Historically first opinion veterinary medicine has focussed on individual animal health giving limited consideration to the benefits of 'herd health' apart from routine vaccination procedures or prophylactics such as worming and parasitic control (Thrusfield, 2013). However this is where clinical epidemiology and evidence-based veterinary medicine converge as summarised by Margaret Slater, "Epidemiology research is concerned with the study of populations and Evidence-based veterinary medicine should be about taking population data and applying it to individual patients" (Slater, 2010).

More recently there has been a move in the right direction and companion animal practitioners are becoming aware of the benefits of population research and the aggregation of patient data to examine the prevalence and incidence of diseases and the aetiology of disease in populations (Faunt et al., 2007, Radford et al., 2010, O'Neill et al., 2012, Thrusfield, 2013, Anholt, 2013).

A critical question is therefore, do veterinarians take research findings and critically appraise the information provided and do they have a full understanding of study design, study limitations and statistical validity (Slater, 2010)? Research to answer this question is currently lacking but this is an interesting question when we consider the research currently offered by popular CPD events such as that presented at BSAVA and reviewed here.

Whilst the strongest evidence (such as systematic reviews and RCTs) may not be available currently, with a few notable exceptions (Olivry and Mueller, 2003, Ferguson et al., 2006, Nuttall and Cole, 2007, Nuttall et al., 2009, Downes et al., 2013), excellent information is available in other forms and presented at more academic or specialised research meetings. Therefore a more important question may be, do vets know the difference between the types of evidence available and have sufficient time to critically appraise the research to make proper use of the information provided?

In 1989 an article was published in the Canadian Veterinary Journal on future directions for veterinary medicine (Unknown., 1989). The report highlights the need for research in veterinary practice and the importance of starting this at an early stage such as undergraduate teaching within veterinary colleges and universities. They also produced a summary of recommended future directions for veterinary medicine and number 4 on the list was to make research a higher priority for individual veterinarians. This is something that has been proposed by others particularly in reference to increased concerns over bio-security (Slater, 2010, Cockcroft and Holmes, 2003). Graduates entering the veterinary profession, along with their more senior colleagues, are increasingly required by clients and regulatory bodies for the veterinary profession and human health to show their treatment decisions are based on valid and robust evidence as an error could have dramatic consequences for both animal health and public health (Cockcroft and Holmes, 2003).

As information technology develops, and a wealth of information is accessible via the internet, clients are becoming more informed and veterinarians are increasingly required to justify their decisions and the relative cost of treatment. Regardless of the move towards EVM in first opinion veterinary practice, veterinarians need evidence to support their treatment decisions and this evidence is currently lacking. The limited opportunities for practice-based research have been discussed here in detail. However to conclude a final point raised by Peter Cockcroft and Mark Holmes (2003) is that “one should remember the main consumers of practice-based research are the veterinarians in practice” and by supporting and engaging in research, such as that presented here, more answers to the many clinical questions faced daily by clinicians will become available, putting research into practice.

The method of information extraction and content analysis of the veterinary EPR presented here allows for a continuation of the EVM cycle introduced in the first chapter of this thesis and presented below. By providing a tool for practice-based research without increasing the work load of the veterinary professional, we can facilitate clinical epidemiology and patient centred research in first opinion practice. The relationship between veterinary clinical epidemiology and evidence-based veterinary medicine is cyclic and ongoing. Veterinary epidemiological studies using data collected directly from the practice computer can provide the necessary evidence needed to support an EVM way of working. Evidence-based veterinary medicine is at the clinically applied end of the spectrum of epidemiology and well-designed veterinary epidemiology studies can provide much needed information to support veterinarians in practice to make clinical decisions using the best evidence available in the form of research results.

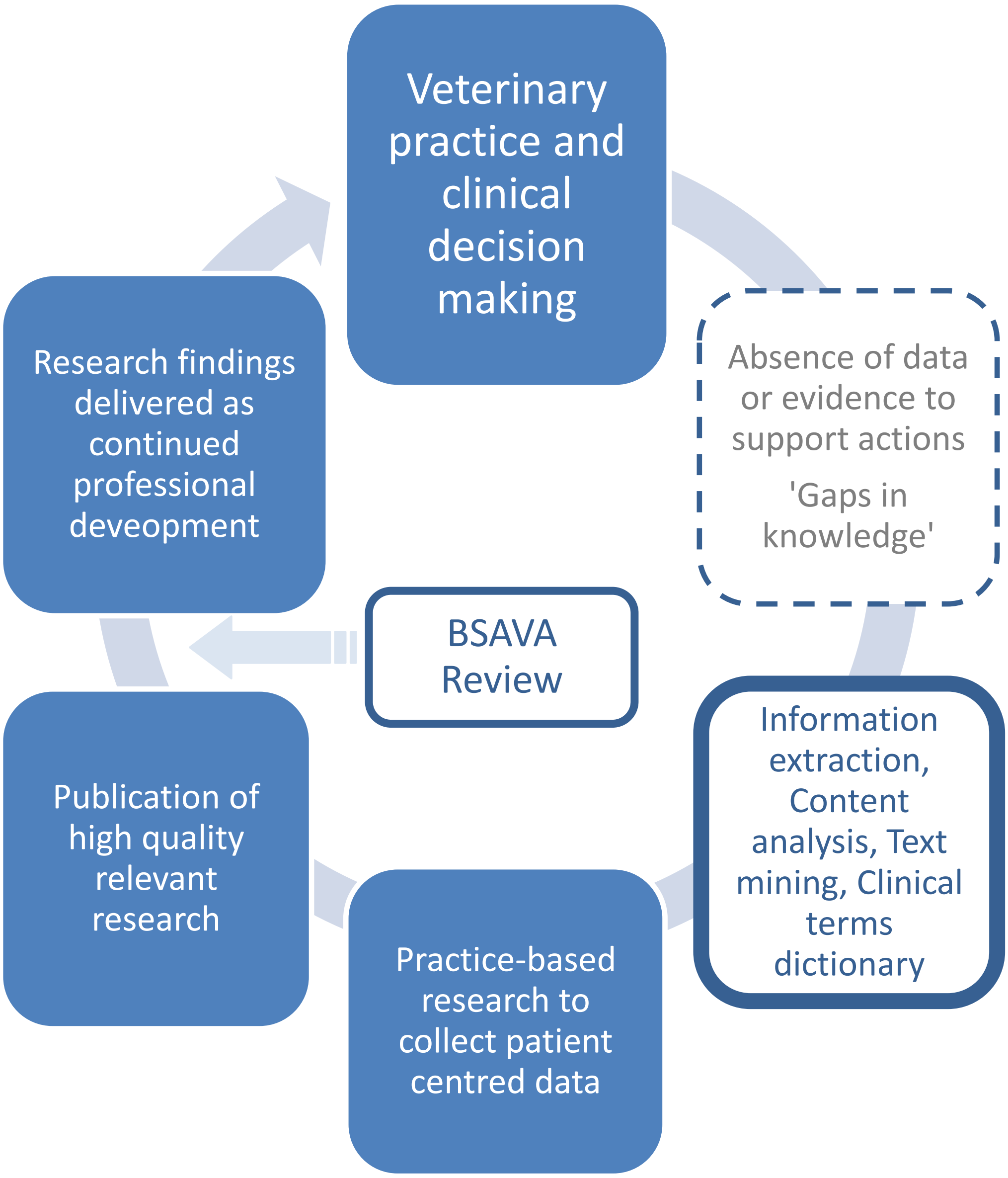


Figure 7-1 The complete cycle of practice-based research and CPD in evidence-based veterinary medicine and clinical decision making.

7.5 Recommendations

This project has produced the only system currently in use in the UK for the extraction and aggregation of patient records from small animal practice which is flexible enough to work across different practice management software systems remotely. In addition the thesis presents a unique method for the analysis of the EPR content using disease specific bespoke clinical terms dictionaries. As such there is a wealth of opportunity for future development.

Recommendations resulting from the findings of this research include the introduction of the XML schema into a second PMS system to validate the method across a second system. Ultimately the aim will be to create a network of sentinel veterinary practices with many different operating systems to combine data and extract information and create a data warehouse of extracted clinical information for use in veterinary clinical epidemiology and practice-based research.

The use of a method of automated extraction is unique within the veterinary field and should be examined next; possibly by collaboration with data communication companies such as Vet Envoy or by applying cloud based data pull rather than push technology for data transfer which would remove the need for the direct involvement of veterinarians in transferring the data. In addition this would allow more time for the CEVM to develop and roll-out a program of training to work more closely with the sentinel veterinarians. In this way the CEVM can ensure the focus of the data collection is as accurate as possible and also deal with queries as they arise. This would also encourage and guide proper use of the system and extraction tool to ensure data integrity.

Research to examine the relationship between the veterinarian and the practice computer/EPR would be beneficial including using qualitative research methods such as focus groups or surveys of the profession. This would allow a fuller understanding of the disparity between the information discussed during the consult and that recorded in the EPR.

The work presented here on the use of clinical dictionaries for content analysis and text mining is both innovative and novel. Until standardised terminology or a universal clinical code is in place across the UK, this may be the only accurate method to text mine the free text within the patient record. Future work will involve developing this method further to include clinical dictionaries for clinical signs and diagnosis associated with animal disease. In addition there may be an opportunity using this methodology to examine how the veterinary profession use clinical terminology within their patient record. It may be that the majority of the profession repeatedly use the same handful of phrases rather than the full list of possible terms. By identifying these key phrases the time taken to text mine patient records could be reduced considerably opening up new opportunities for research and supporting the current work in the field on veterinary terminology standards.

A major hurdle to the veterinarian completing the patient record in full appears to be the time it takes to complete all required fields. By working more closely with PMS system providers and the veterinarians who use them, work could progress to streamline the data entry process to encourage greater capture of information. The active fields could act as a prompt for information and auto-correct could ensure all clinical signs or diagnoses are added to the record using a standardised clinical terminology. Finally it would be valuable to invest more time to further develop the data warehouse. The capacity for storage and management of extracted information needs to increase for future research with many different PMS systems. The increased capacity will ensure accessibility to information and to support data analysis for practice-based research, including longitudinal research such as cohort studies, which are much needed in veterinary practice for outcome research.

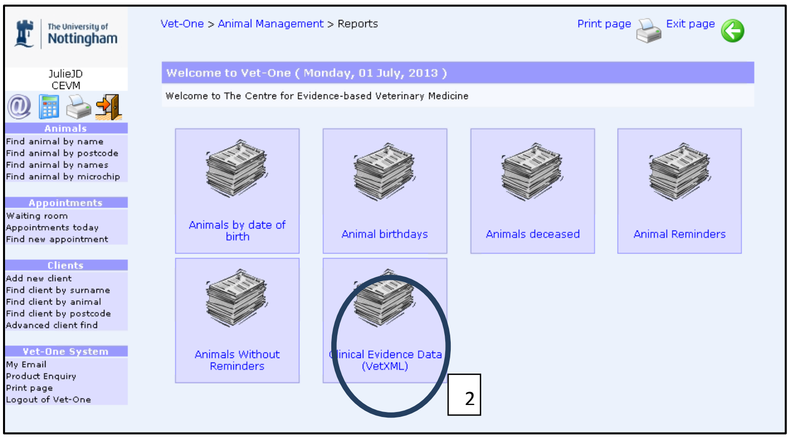
Appendices

Appendix 1 List of practice management systems currently members of the VetXMLConsortium

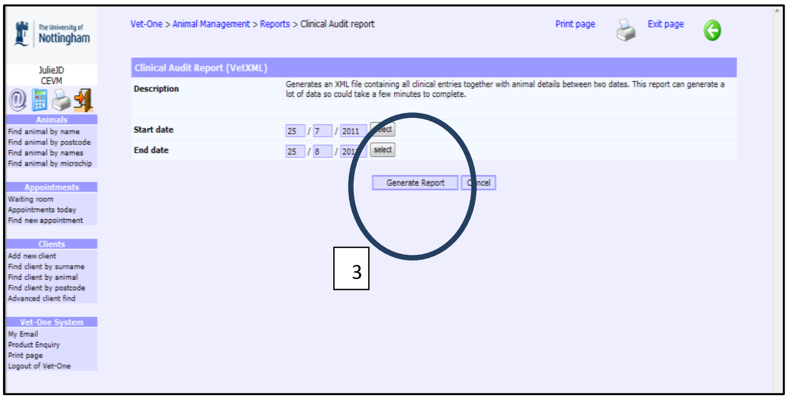
System	**Contact**	**Promotional Details**
AT Veterinary systems	http://www.vetsystems.com	AT has 20 years experience of designing and implementing management solutions for veterinary practices, hospitals, referral centres and universities.
Butler-Schein	http://www.henryschein.com/us-en/Veterinary.aspx http://www.avimark.net/ https://www.impromed.com/	A UK and US based practice management system provider including McAllister Software Systems.
Ez Office Systems (ezVetPro)	http://www.ezofficesystems.com	Ez Office Systems Ltd offers modern advanced veterinary systems to hundreds of veterinary practices in the UK. Comprehensive powerful solutions are accessible through easy to use intuitive screens – engendering loyal and enthusiastic support from existing and new users.
IDEXX Computer Systems	http://www.idexx.com	
Jupiter Systems	www.jupitersystems.net	Specialise in the development and supply of Vet Management Systems. No release date has been given yet for the eClaim service.
Midshire & Ventana	http://www.midshire.com	Midshire Veterinary Systems have been trading since 1991. From day one the company has had one aim – to provide Veterinary Professionals with computer systems designed to complement and enhance their businesses
Rx Works	http://website.rxworks.com/website	Developers of practice management systems, with active practices throughout the World. 1,600 clients in 19 countries
Teleos	http://www.teleosvet.co.uk	Teleos Systems Limited is a specialist company dedicated to providing innovative, cost-effective computing solutions to all types of Veterinary Practice.
Vet-One	http://www.vet-one.co.uk	Vet-One is a new generation of Veterinary Management Software, designed by experienced Veterinary professionals and developed by leading IT analysts. Utilising the latest in secure Internet systems, Vet-One is powerful and yet accessible.
Vet Solutions (PremVet & RoboVet)	http://www.vetsolutions.co.uk	Provide high quality practice management solutions for the Veterinary Profession
Verifac	http://www.verifac.com/	Verifac Veterinary Computers introduced the World's first integrated Veterinary Practice Management Computer System

Appendix 2 Screen shot of CEVM PMS animal management screen and commands used for XML Schema extraction.

1. Reports link,



2. Clinical Evidence Data (VetXML) link,



3. The report request page with the date entry fields and the generate report button.

Appendix 3 Letter to veterinary group for advice on method to opt a client out of the study and to provide information to inform clients for consent

School of Veterinary Medicine and Science
The University of Nottingham
Sutton Bonington Campus
Loughborough
LE12 5RD

FAO: [redacted]

26th April 2012

Dear [redacted]

Please find enclosed two documents for your attention.

I have enclosed another set of laminated posters for informed consent x 4 to be placed around the practice to inform your clients that you are a research practice and the nature of the research.

In addition I have enclosed a number of laminated sheets with instruction on how to opt a client out of the Vet-one data extraction project should they wish to be excluded from data collection during the consultation.

If you have any queries please do not hesitate to contact me.

Warmest regards

Julie Jones-Diette

CENTRE FOR EVIDENCE-BASED VETERINARY MEDICINE

Putting research into practice

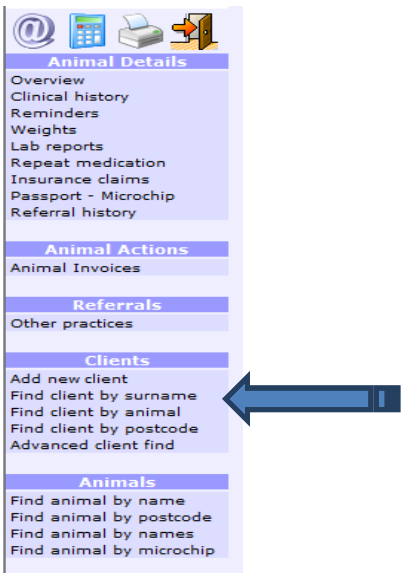
The Centre for Evidence-based Veterinary Medicine (CEVM), based at the University of Nottingham Vet School and Hunters Lodge Veterinary practice are working together to learn more about the diseases that affect pets in the United Kingdom. To do this we need to record why animals are brought to the vets, what is wrong with them and what treatment they are given. This information is very important to us and by working together we can build on our knowledge of the problems pets encounter to improve the health of our pets. Today when you see the vet, the details of the consultation will be recorded, as usual, by the vet using a practice management system on the practice computer. With your permission this data will be used to record what happens during the consultation. We will not ask you or your vet to do anything or answer any questions so your visit to the vets will proceed in the usual way.We will record the species, breed and age of your animal and some information about why you have brought your pet to see the vet today. We will NOT record your name, your pets name or your address. This means any information collected is completely anonymous.

If you do not wish your pet to be included in this study, please tell reception or the vet when you enter the consulting room. If you want to withdraw your pet from the study, you can do this at any time by contacting the practice or the Centre for Evidence-based Veterinary Medicine on 0115 951 6575 or CEVM@nottingham.ac.uk.

Thank you in anticipation of your help with this important study

Instructions to opt a client out from the
CEVM Vet-One VetXML clinical evidence project

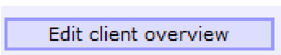
If a client wishes to opt out from the CEVM data capture trial please follow these instructions;



In the Vet-One system from the "**animals history page**" go to the "**clients overview page**".

The **clients overview page** can be reached via any of the client selections on the left hand banner (except add new client) and entering the **client's name.**

8 Once in the **clients overview page**, click on the <**<u>Edit client overview</u>**> button.

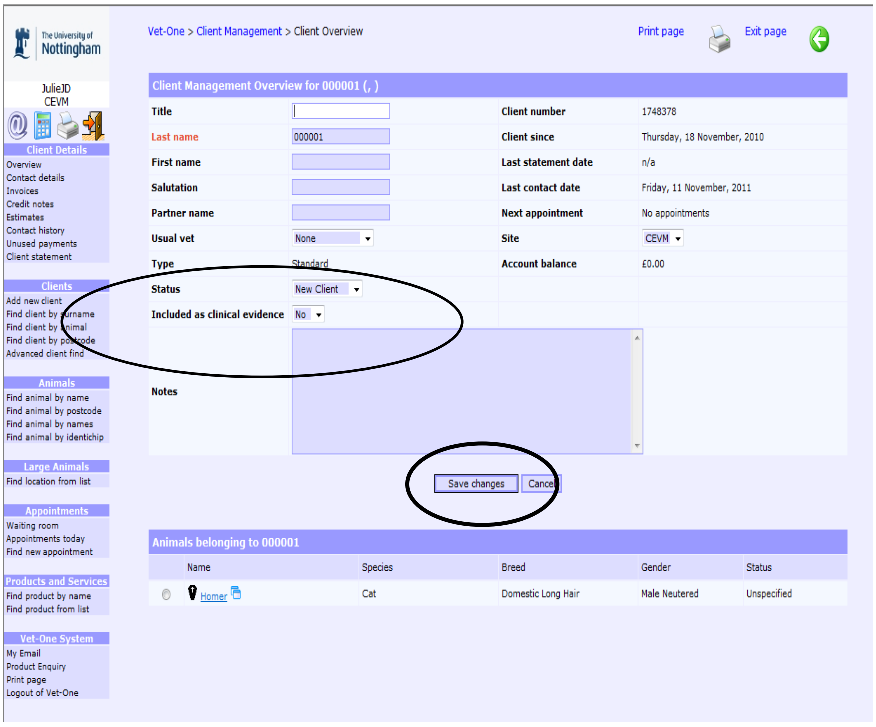


9 Next to the option **"Included as clinical evidence",** use the drop down arrow to change the **<Yes>** to **<No>** and click the **<Save changes>** button.

10 That's it the client is now opted out.

11 No client can return to the trial once they have opted out. Therefore if the client is already set to **NO** in the **Included as clinical evidence** field this should not be altered back to **YES**. This is to ensure the data collection meets the requirements of informed consent and that set out by the ethics review board of The University of Nottingham.

12 No data collected during home visits or farm visits should be included in the data collection.



Appendix 5 Table of characters often found to be in the extracted text from the sentinel PMS system

Termed cleaned from text	No of instances in excel downloads wk 9-16
 	1656
<p>	2768
</p>	3568
	75414
	37
	1469
£	66
 	2659
<table border="0" cellspacing="1" cellpadding="0" width="100%">	13
</td>	426
</tr>	198
<td	429
<tbody>	50
</tbody>	38
<tr class="clientTable">	19
<pre><span style="font-family: courier new,courier;">	69
</table>	38
</pre>	77
<p align="left">	39
	322
&dagger	2
<div class="plainTextContent">	8
	347
‘	1
’	18
“	2
”	2
&	23
<	8
>	8
–	3
Total	89777

i) Search sensitivity (*Se*), a measure of recall provides the proportion of true positives for retrieval success;

$$\frac{\text{True positive (gold standard) cases retrieved}}{\text{All positive cases retrieved}} \times 100$$

ii) *Positive predictive value* (*PPV*), is a measure of precision and calculates the relevance of the search. It is used to measure the proportion of actual cases which are identifiable;

$$\frac{\text{True positive (gold standard) cases retrieved}}{\text{All positive and negative cases retrieved}} \times 100$$

iii) Search specificity (*Sp*), a measure of recall provides the proportion of true negatives for retrieval success;

$$\frac{\text{True negatives (gold standard) cases retrieved}}{\text{All negative cases retrieved}} \times 100$$

iv) *Negative predictive value (NPV*), is a measure of precision and calculates the relevance of the search. It is used to measure the proportion of actual cases which are identifiable;

$$\frac{\text{True negative (gold standard) cases retrieved}}{\text{All positive and negative cases retrieved}} \times 100$$

v) Additional measures included;

Observed prevalence – The proportion of animals in the study that are positive for the condition of interest

Confidence intervals – Gives an indication of the power of the study and the likelihood of detecting a relevant effect. For this study 95% Confidence intervals (CIs) were used.

Appendix 7 List of possible terms used in ER of PMS systems in practice

Reference from: *http://www.bruning.com/terms/acronyms.php* ***An example of some possible acronyms used by veterinarians when recording clinical data in the ER.***

- fracture
A- - inappetant
A/B - antibiotic
AB - antibiotic
abax - abaxial
abc - antibiotics
abd or abdo - abdomen or abdominal
abd n - abdomen palpated and no abnormality detected
abs or ABs - antibiotics
abx - antibiotics
AC - anterior chamber (of the eye)
ADR - ain't doing right
adv - advise or advice
AG - anal glands
AIHA - auto-immune haemolytic anaemia
AON - all others normal
AWE acute wet eczema
ax - axial
B - booster
BAR - bright and responsive, or bright, alert, responsive
BARH - bright, alert, responsive, hydrated
bid - twice daily
BifN - back if necessary
BIOP - been in owner's possession
BIW - before if worried (as in revisit in 1 week or BIW)
bs or BS - blood sample
BT - blood test
B/W - black and white
C - cervical (vertebra - usually followed by a number to denote which one)
C - cough
C+ - moderate cough
C++ - severe cough
C- - no cough
CAH - Chronic Active Hepatitis
caud - caudal
CBA - cat bite abscess
CBC - complete blood count (= routine haematology)
Cc - coccygeal (vertebra - usually followed by a number to denote which one)
CCHF - chronic congestive heart failure
cd - caudal
C7d - recheck in 7 days
CDRM - chronic degenerative radiculomyelopathy
CDS - cat and dog shelter
ce - clinical examination
cep - Ceporex
CHF - congestive heart failure
CHF - chronic heart failure
CKCS - Cavalier King Charles Spaniel
Co - coughv COP - cyclophosphamide/oncovin/prednisolone (chemotherapy protocol)
cr - cranial
CRF - chronic renal failure
CRI - chronic renal insufficiency
CRI - constant rate infusion
CRT - capillary refill time
CS - clinical signs
CS - corticosteroids
C/S or C&S - culture and sensitivity
CSK - Chronic Superficial Keratitis (otherwise known as Pannus)
c-v or cv - cardio-vascular
CVA - cardiovascular accident

Cx - coccygeal (vertebra - usually followed by a number to denote which one)
d - days
D- - diarrhoea absent or diminished
D+ - diarrhoeav D++ - drinking increased
D++ - marked diarrhoea
D+++ - severe diarrhoea
DBA - dog bite abscess
DCM - dilated cardiomyopathy
DD, DDx or DDxx - differential diagnosis
D&D - destroy and disposal
DHLPPi - distemper, hepatitis, Leptospirosis, parvovirus and parainfluenza: dog vaccine covering these diseases
DI - diabetes insipidus
DIC - disseminated intravascular coagulation
DJD - Degenerative Joint Disease
DLH - domestic long hair (non-pedigree cat)
DM - diabetes mellitus
do - days old (following a number)
D/O - drain out (removal of surgical drain)
D/O - dropped off (for treatment or examination)
DOA - dead on arrival
dors - dorsal
DSH - domestic short hair (non-pedigree cat)
Dx - diagnosis
E&A - euthanasia and aftercare
EAG - empty (or express) anal glands
EBT - English Bull Terrier
E/D+ - eating and drinking normally
EDUD - eating, drinking, urinating and defaecating
EOD - every other day
EPI - exocrine pancreatic insufficiency
est - estimate
e/t - endo-tracheal
EUA - examine under anaesthetic
euth - euthanase
ex or ext - external
e(x)d - every (x) days
ex lap - exploratory laparotomy
e(x)wks - every (x) weeks
F - female (see also Fe and Fn)
F1/4 - forequarters
FAD - flea allergic dermatitis
fat++ - obese
FB - foreign body
Fe - female (entire)
FeLV - feline leukaemia virus
FIA - feline infectious anaemia (Haemobartonellosis)
Fin - finish this treatment
FIP - feline infectious peritonitis
FIV - feline immunosuppressive virus
f/l or FL - front legs
FLUTD - feline lower urinary tract disease
FMD - feline miliary dermatitis
Fn - female (neutered)
FNA - fine needle aspiratev FNAB - fine needle aspirate biopsy
FQ - fore quarters
FUO - fever of unknown origin
FUS - feline urological syndrome
ga or GA - general anaesthetic
GDV - gastric dilitation and volvulus
GORK - God only really knows
GSD - German Shepherd Dog
GSHP - German Short-Haired Pointer
GSP - German Short-haired Pointer
H1/4 - hindquarters
H+ - history
HAC - hyperadrenocorticism

HBB - hit by bullet
HBC - hit by car
HBT - hit by truck
HBTr - hit by train
HCM - hypertrophic cardiomyopathy
HCT - haematocrit
HD - hip dysplasia
HGE - haemorrhagic gastro-enteritis
h/l or HL - hind legs
HQ - hind quarters
HR - heart rate
hrs - hours
HRS - high rise syndrome (animal fell off balcony/windowsill)
H's - Horner's syndrome
HX or Hx - history
HyperT4 - hyperthyroidism
HypoT4 - hypothyroidism
I - incisor (tooth - followed by a number to indicate which one)
ID - identification or Identichip
i/d - inter-digital
iid - once a day
im or i/m - intramuscular
IMHA - immune-mediated haemolytic anaemia
impr - improved: followed by a variable number of +'s to indicate degree of improvement
IMTP - immune-mediated thrombocytopenia
inj - inject or injection
int - internal
IOL - Intra Ocular Lens
IOP - intra-ocular pressure
ip or i/p - intra-peritoneal
iv or i/v - intravenous
JAR - "just ain't right"
jt - joint
KC - kennel cough (infectious tracheobronchitis)
KCS - Kerato Conjunctivitis Sicca (otherwise known as Dry Eye)
L - lumbar (vertebra - usually followed by a number to denote which one)
L - left
LA - left atrium
LA - local anaesthetic
LAE - left atrial enlargement
lat - lateral
BD/LD - big dog/little dog (re fight injuries)
LF - left fore
LH - left hind
LHD - long haired domestic (non-pedigree cat)
LI - large intestine
LIU - Lens Induced Uveitis
LL - left lower (eg. eyelid, lip)
ln - lymph node/s
LTF - learning to fly (animal fell off balcony or similar)
LU - left upper (eg. eyelid, lip)
LV - left ventricle
LVE - left ventricular enlargement
M - molar (tooth - followed by a number to indicate which one)
M - male (see also Me and Mn)
manip - manipulation
m/c - metacarpal
MD - myocardial disease
Me - male (entire)
med - medial
min - minutes
mm or Mmb - mucous membranes
Mn - male (neutered)
mo - months old (following a number)
MT - mammary tumour
m/t - metatarsal
MVA - motor vehicle accident
N - normal
NAD - no abnormality detected (or detectable)

NAF - no abnormal findings
nbg - treatment not working
n def - nerve deficit
NNTSA - no need to see again
NOAD - no other abnormality diagnosed
NOAF - no other abnormal findings
no C/S/V/D - no coughing, sneezing, vomiting or diarrhoea
NOS - no other symptoms
NPO - Nothing per os/nil per os
N/S - next step
NSAID - non-steroidal anti-inflammatory drug
NSF - no significant finding
nyd - not yet diagnosed
O - owner
OA - osteoarthritis
OCD - osteochondritis dissecans
OCNE - off colour, not eating
OES - Old English Sheepdog
O/F - off food
OOH - out of hours
op - operate or operation
P - pulse
palp - palpate or palpation
PAN - peri-arteritis nodosa
PBT - Pit Bull Terrier
pcv - packed cell volume (haematocrit)
PDA - patent ductus arteriosis
PE - physical examination
PPi - parvovirus and parainfluenza: dog vaccine covering these diseases
PPP - pretty poor prognosis
PM - pre-molar (tooth - followed by a number to indicate which one)
PM - post mortem
PTS - put to sleep (euthanase)
PD - polydipsia
pop - popliteal (lymph node)
pred - prednisolone
prn - as needed (pro re nata - Latin)
prob - probably or problem
PU/PD - polyuria/polydipsia
PUO - pyrexia of unknown origin
Px - physical examination or physical examination findings
Px - prognosis
q200 - owner has been given an estimate of £200 for this procedure (q for quoted)
qid - four times daily
QOL - quality of life
R - right
RA - right atrium
RAE - right atrial enlargement
rbc - red blood cell
RC 1w - recheck in 1 week
RCM - restrictive cardiomyopathy
RDVM - referring Dr Veterinary Medicine
Rex - re-examine
RF - right fore
RH - right hind
RI - re-examine/re-inspect
RI 7d - return in 7 days
RL - right lower (eg. eyelid, lip)
rMRCVS - Referring MRCVS
R/O - rule out
rost - rostral
rpt - repeat
rrm - restricted range of movement, eg in a joint
rrr - resting respiration rate
RS - remove sutures
RTA - road traffic accident
RU - right upper (eg. eyelid, lip)
rv - recheck
RV - right ventricle
RVE - right ventricular enlargement
Rx - prescription or dispense
Rx - re-examine
sc or s/c - subcutaneous/subcuticular

SBI - something bad inside, diagnosis of exclusion for ADR
SBT - Staffordshire Bull Terrier
SHD - short haired domestic (non-pedigree cat)
SI - small intestine
SI - self inflicted
SIBO - small intestinal bacterial overgrowth
sid - once daily
sig - signalment
SIT - self-inflicted trauma
s/m - submandibular
Sn - sneezing
S/O - stitches (or staples) out
S&P - scale and polish (of teeth)
sq - subcutaneous
S/R - suture removal
SS - skin scrape
STO - speak to owner
SWB - salt water bathing
Sx - surgery
T - thoracic (vertebra - usually followed by a number to denote which one)
Tb - tabby
Tb/W - tabby and white
TGO - to go home
tid - three times daily
T-L - thoraco-lumbar
TLI - trypsin-like immunoreactivity
TMJ - temporo-mandibular joint
Tn or TN - temperature normal
TNT - toenail trim
Tort - tortoiseshell
Tort/W - tortoiseshell and white
TPN - total parenteral nutrition
TPR - temperature, pulse and respiration
TPRn - temperature, pulse and respiration all normal
TTO - talked to owner
TWBC - total white blood cells
twds - towards
Tx - treatment
U/A or UA - urinalysis
UGA - under general anaesthetic
ULA - under local anaesthetic
URT - upper respiratory tract
URTI - upper respiratory tract infection
US - urine sample
US or us - ultrasound
UTI - urinary tract infection
V- - vomiting absent or diminished
V+ - vomiting
V++ - moderate vomiting
V+++ - severe vomiting
vacc - vaccinate or vaccine
V&D or V/D+ - vomiting and diarrhoea
V/D - : no vomiting, no diarrhoea
VMB - very much better
vmi - very much improvedv wbc - white blood cell
wmd - white meat diet
WNL - within normal limits
wo - weeks old (following a number)
wt - weight
wt!! - overweight
x/60 - x minutes
x/24 - x hours
x/7 - x days
x/52 - x weeks
x/12 - x months
xr or XR - x-ray
yo - years old

Appendix 8 WordStat output for text mining of sentinel EPR using a dictionary of term.

Dictionary 1.0 Positive terms identified are in bold.

1. [Case #1 NS1_ANIMAL = 59 VARIABLE = NS1_TEXTEN]

   Showing signs of cystitis but now resolved. **Bladder** small but non painful. ADv urine sample to check for crystals. Given katkor. Put onto urinary.

2. [Case #1 NS1_ANIMAL = 59 VARIABLE = NS1_TEXTEN]

   Showing signs of **cystitis** but now resolved. Bladder small but non painful. ADv urine sample to check for crystals. Given katkor. Put onto urinary.

3. [Case #1 NS1_ANIMAL = 59 VARIABLE = NS1_TEXTEN]

   Showing signs of cystitis but now resolved. Bladder small but non painful. ADv urine sample to check for crystals. Given katkor. Put onto **urinary**.

4. [Case #1 NS1_ANIMAL = 59 VARIABLE = NS1_TEXTEN]

   Showing signs of cystitis but now resolved. Bladder small but non painful. ADv **urine** sample to check for crystals. Given katkor. Put onto urinary.

5. [Case #36 NS1_ANIMAL = 4085 VARIABLE = NS1_TEXTEN]

   wt loss missing 3d not eating today not drinking much 38.2 nad otherwise **bladder** full abdomen ok bs lab

6. [Case #58 NS1_ANIMAL = 5773 VARIABLE = NS1_TEXTEN]

   Not eaten much over past week to 10 days. heart rate 190 bpm. Has been showing signs of **cystitis**. No vomiting. HAs dought cranial abdomen. Suspect something going on in stomach or liver. Try b12 nad if no imp a week then adv checking T4 levels.

7. [Case #62 NS1_ANIMAL = 5956 VARIABLE = NS1_TEXTEN]

Bloods show severe anaemia. Mass in cr abdomen. Suspect bleeding into gut. **Bladder** 3 x normal size. PTS.

8. [Case #129 NS1_ANIMAL = 12410 VARIABLE = NS1_TEXTEN]

   Insurance claim sent to Direct Line for Continuation of **Dysuria** 03/12/11 - 03/02/12 invoices E/67503 & 68138 115.72 payable to HL

9. [Case #130 NS1_ANIMAL = 12410 VARIABLE = NS1_TEXTEN]

   Bladder scan shows 12-20 1-2mm opacities in **bladder**, v small stones no sign of larger stone noted prev. U+ sample diff to get as bladder small pH = 6.5 so ok. Adv try on s/d as will reduce pH further, rechekc before end of bag 3-5wks.

10. [Case #130 NS1_ANIMAL = 12410 VARIABLE = NS1_TEXTEN]

    Bladder scan shows 12-20 1-2mm opacities in bladder, v small stones no sign of larger stone noted prev. U+ sample diff to get as **bladder** small pH = 6.5 so ok. Adv try on s/d as will reduce pH further, rechekc before end of bag 3-5wks.

11. [Case #130 NS1_ANIMAL = 12410 VARIABLE = NS1_TEXTEN]

Bladder scan shows 12-20 1-2mm opacities in bladder, v small stones no sign of larger stone noted prev. U+ sample diff to get as bladder small pH = 6.5 so ok. Adv try on s/d as will reduce pH further, rechekc before end of bag 3-5wks.

12. [Case #169 NS1_ANIMAL = 13514 VARIABLE = NS1_TEXTEN]

overnight some blood in urine thirst appetite normal t38.1 fed purina low pr + denes **bladder** ok highly strung req us start metacam

13. [Case #169 NS1_ANIMAL = 13514 VARIABLE = NS1_DIAGNO]

Cystitis

14. [Case #169 NS1_ANIMAL = 13514 VARIABLE = NS1_TEXTEN]

overnight some blood in **urine** thirst appetite normal t38.1 fed purina low pr + denes bladder ok highly strung req us start metacam

15. [Case #263 NS1_ANIMAL = 17535 VARIABLE = NS1_TEXTEN]

GA and repaired large diaphragmantic hernia. Colon SI and liver into chest. No serious damage to these organs. **Bladder** intact. Closed diaphragm with 2 layer pds. Post op rads v good. Good inflation of lung fields.Small amount of fluid ventro caudal lung fields but not enough to warrant chest drain. Further radiographs taken of right hind. Fracture of all four metatarsals. Also fractureT 13.May need stabilisation. Send to FP's. Has movement and deep pain of hindlimbs. Is urinating normally.

16. [Case #283 NS1_ANIMAL = 17583 VARIABLE = NS1_TEXTEN]

2nd op for CRF tx pet doctors. O concerned U+ bloods nor K+ assessed recently. HR 110, O having to express **bladder** (7d) and bowels (24hs) v weak on HLs gen ataxia, did eat 1 tsp this am and brighter since defaecated last night. Adv tx can be continued and furthered BP, K+ anti emetics etc. at present cat severely dislikes mirtazepine, and other meds are also abattle. Adv any small improvements to mm weakness may enable another few weeks of tx perhaps. O admitted that tx struggle and cat had poor QOL last week, so despite tx could be taken further more likely to regret keeping him going too long rather than pts too soon. Not too soon in this case as many owners would have pts sooner, these owners v dedicated butas Px poor putting this cat who hates meds through further tx is difficult to justify.

Dictionary 1.1 Positive terms identified are in bold.

1. [Case #1 NS1_ANIMAL = 59 VARIABLE = NS1_TEXTEN]

   Showing signs of cystitis but now resolved. **Bladder** small but non painful. ADv urine sample to check for crystals. Given katkor. Put onto urinary.

2. [Case #1 NS1_ANIMAL = 59 VARIABLE = NS1_TEXTEN]

   Showing signs of **cystitis** but now resolved. Bladder small but non painful. ADv urine sample to check for crystals. Given katkor. Put onto urinary.

3. [Case #1 NS1_ANIMAL = 59 VARIABLE = NS1_TEXTEN]

   Showing signs of cystitis but now resolved. Bladder small but non painful. ADv urine sample to check for crystals. Given katkor. Put onto **urinary**.

4. [Case #1 NS1_ANIMAL = 59 VARIABLE = NS1_TEXTEN]

   Showing signs of cystitis but now resolved. Bladder small but non painful. ADv **urine** sample to check for crystals. Given katkor. Put onto urinary.

5. [Case #16 NS1_ANIMAL = 1484 VARIABLE = NS1_TEXTEN]

   Sedate midz/ket, xray NAD chest or abdo. **ultrasound** NAD abdo stomach WNL. No reason for V+. eaten mod amount in hosp today. v small goitre. with t4 at top end of normal suggest trial tx with vidalta. c7d. can eat without apparant discomfort. mild constipation. F+ when recovered. STT1 >15 after 30s.

6. [Case #30 NS1_ANIMAL = 3294 VARIABLE = NS1_TEXTEN]

Wt finally increasing eating well now. Adv cont to monitor weight and **feliway** 1 more month.

7. [Case #36 NS1_ANIMAL = 4085 VARIABLE = NS1_TEXTEN]

   wt loss missing 3d not eating today not drinking much 38.2 nad otherwise **bladder** full abdomen ok bs lab

8. [Case #58 NS1_ANIMAL = 5773 VARIABLE = NS1_TEXTEN]

   Not eaten much over past week to 10 days. heart rate 190 bpm. Has been showing signs of **cystitis**. No vomiting. HAs dought cranial abdomen. Suspect something going on in stomach or liver. Try b12 nad if no imp a week then adv checking T4 levels.

9. [Case #62 NS1_ANIMAL = 5956 VARIABLE = NS1_TEXTEN]

Bloods show severe anaemia. Mass in cr abdomen. Suspect bleeding into gut. **Bladder** 3 x normal size. PTS.

10. [Case #68 NS1_ANIMAL = 6187 VARIABLE = NS1_TEXTEN]

Stillu / f incontinence 1-2 / wk. Nervous disposition Not using litter tray regularly and then only u not f Is urinating on anything soft, cushions, sofa. Really sounds behavioural. Scratches furniture- mostly indoor cat O thinks things improve with **feliway**. Did not do this when other cat was around. Maybe he is feeling insecure being on his own. crepitus rh hip ^thirst Temp 99F. Likley needs to see behaviourist. occasional v 1-2 / months shortly after eating rubs up on furniture / people a lot bs lab overdue vaccination restart course nob tri a280a01 felv 3dkv02

11. [Case #68 NS1_ANIMAL = 6187 VARIABLE = NS1_TEXTEN]

Stillu / f incontinence 1-2 / wk. Nervous disposition Not using litter tray regularly and then only u not f Is **urinating** on anything soft, cushions, sofa. Really sounds behavioural. Scratches furniture- mostly indoor cat O thinks things improve with feliway. Did not do this when other cat was around. Maybe he is feeling insecure being on his own. crepitus rh hip ^thirst Temp 99F. Likley needs to see behaviourist. occasional v 1-2 / months shortly after eating rubs up on furniture / people a lot bs lab overdue vaccination restart course nob tri a280a01 felv 3dkv02

12. [Case #95 NS1_ANIMAL = 9169 VARIABLE = NS1_TEXTEN]

wt loss cont eating better but not vast quantaties. bg 6.3 - good. Adv monitor 1wk, try a/d in mean time to boost energy. o to monitor for **pupd**.

13. [Case #129 NS1_ANIMAL = 12410 VARIABLE = NS1_TEXTEN]

Insurance claim sent to Direct Line for Continuation of **Dysuria** 03/12/11 - 03/02/12 invoices E/67503 & 68138 115.72 payable to HL

14. [Case #130 NS1_ANIMAL = 12410 VARIABLE = NS1_TEXTEN]

Bladder scan shows 12-20 1-2mm opacities in bladder, v small stones no sign of larger stone noted prev. U+ sample diff to get as bladder small pH = 6.5 so ok. Adv try on s/d as will reduce pH further, rechekc before end of bag 3-5wks.

15. [Case #130 NS1_ANIMAL = 12410 VARIABLE = NS1_TEXTEN]

Bladder scan shows 12-20 1-2mm opacities in **bladder**, v small stones no sign of larger stone noted prev. U+ sample diff to get as bladder small pH = 6.5 so ok. Adv try on s/d as will reduce pH further, rechekc before end of bag 3-5wks.

16. [Case #130 NS1_ANIMAL = 12410 VARIABLE = NS1_TEXTEN]

Bladder scan shows 12-20 1-2mm opacities in bladder, v small stones no sign of larger stone noted prev. U+ sample diff to get as **bladder** small pH = 6.5 so ok. Adv try on s/d as will reduce pH further, rechekc before end of bag 3-5wks.

17. [Case #130 NS1_ANIMAL = 12410 VARIABLE = NS1_TEXTEN]

Bladder scan shows 12-20 1-2mm opacities in bladder, v small stones no sign of larger stone noted prev. U+ sample diff to get as bladder small **pH** = 6.5 so ok. Adv try on s/d as will reduce pH further, rechekc before end of bag 3-5wks.

18. [Case #130 NS1_ANIMAL = 12410 VARIABLE = NS1_TEXTEN]

Bladder scan shows 12-20 1-2mm opacities in bladder, v small stones no sign of larger stone noted prev. U+ sample diff to get as bladder small pH = 6.5 so ok. Adv try on s/d as will reduce **pH** further, rechekc before end of bag 3-5wks.

19. [Case #130 NS1_ANIMAL = 12410 VARIABLE = NS1_TEXTEN]

Bladder scan shows 12-20 1-2mm opacities in bladder, v small stones no sign of larger **stone** noted prev. U+ sample diff to get as bladder small pH = 6.5 so ok. Adv try on s/d as will reduce pH further, rechekc before end of bag 3-5wks.

20. [Case #150 NS1_ANIMAL = 12899 VARIABLE = NS1_TEXTEN]

Admit last night. Markedly **PUPD** since home breathing stable and normal. no chest sounds. Not eaten very well in last 2d. Adv check bloods incase dehydrated/renal function compromised with diuretics. Urea was 34 creat WNL. Mild neutrophillia also. Adv combination of DM diuresis (not not been giving insulin as not eating at home) and Frusemide (although also not given much at home) has produced azotaemia. hopefully ARF as creat WNL should resolve with IVFT. renal function good previous to this. Adv will recheck Urea Creat sat am. hopefully resolving and can be discharged INI then pts. O concerned as to QOL at present. Adv DM and asthema prev stable so should cont to be so and if ARF resolves the Px good. If tipped into CRF then px worse and pts would be sensible option. Cat has a Hx of stress anorexia so if azotaemia resolves and still anorexic give 2d at home incase will start eating then. Recheck Mon Urea also.

21. [Case #169 NS1_ANIMAL = 13514 VARIABLE = NS1_TEXTEN]

overnight some blood in urine thirst appetite normal t38.1 fed purina low pr + denes **bladder** ok highly strung req us start metacam

22. [Case #169 NS1_ANIMAL = 13514 VARIABLE = NS1_DIAGNO]

Cystitis

23. [Case #169 NS1_ANIMAL = 13514 VARIABLE = NS1_TEXTEN]

overnight some blood in **urine** thirst appetite normal t38.1 fed purina low pr + denes bladder ok highly strung req us start metacam

24. [Case #172 NS1_ANIMAL = 13732 VARIABLE = NS1_TEXTEN]

Straining to pass motions and still lots of blood. O decided to call it a day as cat becoming distressed.

25. [Case #222 NS1_ANIMAL = 16679 VARIABLE = NS1_TEXTEN]

tearing out fur for 6wks or so started over thorax now strip caudal dorsum. ventrum ok. O seen cat removing fur. regular flea tx long term good compliance. O instigated **feliway** 4wks ago no improvement. has flap and goes out. fleas and dirt -ve on comb. Dx ?FAD from 1-2 bites or stress response. given depo-medrone monitor for response. warned will take 6-8wks for fur to regrow.

26. [Case #234 NS1_ANIMAL = 17321 VARIABLE = NS1_TEXTEN]
    6 month check treated advocate all ok continually sucks things for comfort discuss **feliway** lac can
27. [Case #253 NS1_ANIMAL = 17511 VARIABLE = NS1_TEXTEN]

URINALYSIS REPORT 16/01/12 GLU = Neg KET = Neg BLD = Neg **pH** = 6.0 PRO = + LEU = Neg Specific gravity UG = 1.011 (cat >1.025)

28. [Case #253 NS1_ANIMAL = 17511 VARIABLE = NS1_TEXTEN]

URINALYSIS REPORT 16/01/12 GLU = Neg KET = Neg BLD = Neg pH = 6.0 PRO = + LEU = Neg Specific gravity UG = 1.011 (cat >1.025)

29. [Case #261 NS1_ANIMAL = 17535 VARIABLE = NS1_TEXTEN]

Referral letter from Fitzpatrick 30th January 2012 I am sorry I missed you on Saturday and am sorry that you found out the sad news about Ruby from Mrs Raynes.You had done a really thorough workup and made some very good decisions before referring Ruby to us. She was stable on admittance. We radiographed her thorax and there was only a small amount of free fluid present at that time, which was confirmed on MRI. We went straight on to perform surgery to repair the spinal fracture with a plan to repair the metatarsal fractures at a later date. Surgery had gone very well and the anaesthesia was uneventful. Radiographs taken post op showed anatomic fracture reduction and normal spinal alignment. The thorax and abdomen remained as before, with mild pleural effusion only and normal lung patterns. The respiratory arrest happened very suddenly during the night early on Friday morning. The overnight team attempted to resuscitae Ruby but were unable to to do so. Mrs Raynes was informed of the sad news later on Friday morning. We did not perform post mortem exam but the main differential diagnosis are pulmonary thromboembolism and and re inflation injury (fulminant pulmonary oedema). Both are potential complications after an initially successful surgery to repair a diaphragmatic rupture. HISTORY 1 year old cat in owners possession sice a kitten. Assumed RTA at midnight Tuesday Wednesday presentation 3pm Thursday Dissappeared during the day and made it through cat flap. Subsequent treatments at the primary care practice. Bloods showed very high ALT and CK (PCV ok) Workup showed diapghrgmatic rupture Repaired on day of presentation. Apparent ability to **urinate** still present. Radiographs show T13 fracture with 50% step T13-L1 also disatl 1/3 comminuted fractures MT 2-5. EXAM Quiet but responsive CNS exam normal 3rd eyelid protruding (post GA) FL motor function normal HL motor function (careful support) absnt. HL nociception present. Tail movement and sensation present. No attempt to perform full neurological or orthopaedic exams due to known unstable spinal fracture. ASSESSMENT Grade 3 TL myelopathy Unstable T12-13 fracture luxation Unstable right MT fractures MRI MRI confirmed the T12-13 spinal fracture luxation The spinal cord was minimally compressed. SURGERY REPORT Right dorsolateral approach to T12-13 Mini-hemilaminectomy 4pin internal fixator (1.1mm mini interface positive profile pins and bone cement connecting bar Routine closure Mike Farrell BvetMed, CertVA, CertSAS, DipECVS, MRCVS

30. [Case #263 NS1_ANIMAL = 17535 VARIABLE = NS1_TEXTEN]

GA and repaired large diaphragmantic hernia. Colon SI and liver into chest. No serious damage to these organs. **Bladder** intact. Closed diaphragm with 2 layer pds. Post op rads v good. Good inflation of lung fields.Small amount of fluid ventro caudal lung fields but not enough to warrant chest drain. Further radiographs taken of right hind. Fracture of all four metatarsals. Also fractureT 13.May need stabilisation. Send to FP's. Has movement and deep pain of hindlimbs. Is urinating normally.

31. [Case #263 NS1_ANIMAL = 17535 VARIABLE = NS1_TEXTEN]

GA and repaired large diaphragmantic hernia. Colon SI and liver into chest. No serious damage to these organs. Bladder intact. Closed diaphragm with 2 layer pds. Post op rads v good. Good inflation of lung fields.Small amount of fluid ventro caudal lung fields but not enough to warrant chest drain. Further radiographs taken of right hind. Fracture of all four metatarsals. Also fractureT 13.May need stabilisation. Send to FP's. Has movement and deep pain of hindlimbs. Is **urinating** normally.

32. [Case #270 NS1_ANIMAL = 17546 VARIABLE = NS1_TEXTEN]

squatting **straining** to pass solids normal stool produced t= nad on exam some firm pellets in bowel

33. [Case #283 NS1_ANIMAL = 17583 VARIABLE = NS1_TEXTEN]

2nd op for CRF tx pet doctors. O concerned U+ bloods nor K+ assessed recently. HR 110, O having to express **bladder** (7d) and bowels (24hs) v weak on HLs gen ataxia, did eat 1 tsp this am and brighter since defaecated last night. Adv tx can be continued and furthered BP, K+ anti emetics etc. at present cat severely dislikes mirtazepine, and other meds are also abattle. Adv any small improvements to mm weakness may enable another few weeks of tx perhaps. O admitted that tx struggle and cat had poor QOL last week, so despite tx could be taken further more likely to regret keeping him going too long rather than pts too soon. Not too soon in this case as many owners would have pts sooner, these owners v dedicated butas Px poor putting this cat who hates meds through further tx is difficult to justify.

Dictionary 1.2 Positive terms identified are in bold.

1. [Case #1 NS1_ANIMAL = 59 VARIABLE = NS1_TEXTEN]

   Showing signs of **cystitis** but now resolved. Bladder small but non painful. ADv urine sample to check for crystals. Given katkor. Put onto urinary.

2. [Case #1 NS1_ANIMAL = 59 VARIABLE = NS1_TEXTEN]

   Showing signs of cystitis but now resolved. Bladder small but non painful. ADv urine sample to check for crystals. Given katkor. Put onto **urinary**.

3. [Case #1 NS1_ANIMAL = 59 VARIABLE = NS1_TEXTEN]

   Showing signs of cystitis but now resolved. Bladder small but non painful. ADv **urine** sample to check for crystals. Given katkor. Put onto urinary.

4. [Case #58 NS1_ANIMAL = 5773 VARIABLE = NS1_TEXTEN]

   Not eaten much over past week to 10 days. heart rate 190 bpm. Has been showing signs of **cystitis**. No vomiting. HAs dought cranial abdomen. Suspect something going on in stomach or liver. Try b12 nad if no imp a week then adv checking T4 levels.

5. [Case #129 NS1_ANIMAL = 12410 VARIABLE = NS1_TEXTEN]

   Insurance claim sent to Direct Line for Continuation of **Dysuria** 03/12/11 - 03/02/12 invoices E/67503 & 68138 115.72 payable to HL

6. [Case #169 NS1_ANIMAL = 13514 VARIABLE = NS1_DIAGNO]

Cystitis

7. [Case #169 NS1_ANIMAL = 13514 VARIABLE = NS1_TEXTEN]

   overnight some blood in **urine** thirst appetite normal t38.1 fed purina low pr + denes bladder ok highly strung req us start metacam

Appendix 9 Kappa analysis from Minitab 16 (v16.2.2) & list of how Kappa might be interpreted (Landis and Koch, 1977)

Kappa Analysis

Attribute Agreement Analysis for Observation by NR, Observation by RD

Between Appraisers

```
Assessment Agreement

# Inspected  # Matched  Percent       95% CI
         36         30    83.33  (67.19, 93.63)

# Matched: All appraisers' assessments agree with each other.


Fleiss' Kappa Statistics

Response     Kappa  SE Kappa        Z  P(vs > 0)
1          0.66667  0.166667  4.00000     0.0000
2          0.70370  0.166667  4.22222     0.0000
3          1.00000  0.166667  6.00000     0.0000
4          0.65217  0.166667  3.91304     0.0000

Overall    0.72254  0.124695  5.79449     0.0000
```

Kappa	Interpretation
< 0	Poor agreement
0.0 – 0.20	Slight agreement
0.21 – 0.40	Fair agreement
0.41 – 0.60	Moderate agreement
0.61 – 0.80	Substantial agreement
0.81 – 1.00	Almost perfect agreement

Appendix 10 Data worksheet of all consultation data collected by both methods of data collection

Date	N	Observation	Order of presentation by client	Brief description	Extraction (Vet 1 assessed)	Brief description	Extraction (Vet 2 assessed)	brief description	Outcome
14/02/2012	1	3	1	Blepharospasm	2	Conjunctivitis hx lens replacement	1	mild conjunctivitis r eye	ther tx
			2	Diabetic - Bloods awaiting results		Blood reported - fructosamine (diabetes?)			manage/ ther tx
			3	review of prev prob - hx lens sx ok at minute					nothing
14/02/2012	2	1	1	Dental	1	Dental post op check	1	Post op check	nothing
14/02/2012	3	3	1	worm/flea/vacc booster	2	worm/flea/vacc booster	1	worm/flea/vacc booster KC	prop tx
			2	clip nails		clip nails			nothing
			3	Seasonal abnormality had contraceptive jab - not back in season yet					0
14/02/2012	4	2	1	worm/flea/vacc booster	2	worm/flea/vacc booster	1	worm/flea/vacc booster	prop tx

Date	N	Observation	Order of presentation by client	Brief description	Extraction (Vet 1 assessed)	Brief description	Extraction (Vet 2 assessed)	brief description	Outcome
*			2	clip nails		clip nails			prop tx
14/02/2012	5	1	1	Repeat consult for ears	1	Repeat consult for ears	1	Repeat consult for ears	manage/ ther tx
14/02/2012	6	5	1	worm/flea/vacc booster	2	worm/flea/vacc booster	1	worm/flea/vacc booster	prop tx
***			2	weightgain		overweight			manage
*			3	new problem - licking feet, nail penetrating pad					ther tx
*			4	previous problem licking feet					nothing
			5	halitosis - All fine					manage
14/02/2012	7	3	1	post op check of wound	3	post op check of wound	3	post op check of wound. OK inflammation gone ***	ther tx
			2	head tilt		head tilt		head tilt	ther tx
			3	Review of inflamed/swollen joints/carpus		inflammation		inflammation	ther tx

Date	N	Observation	Order of presentation by client	Brief description	Extraction (Vet 1 assessed)	Brief description	Extraction (Vet 2 assessed)	brief description	Outcome
14/02/2012	8	1	1	Review of bloods/ medication review	1	Review of bloods/ medication review	1	Review of bloods/ medication review	ther tx/work up
15/02/2012	1	4	1	vacc/worm	1	vacc/worm	1	vacc	
			2	microchip placement query - no scanner available					nothing
			3	clip nails					nothing
			4	behaviour problem - acting strange avoidance and hiding					nothing
15/02/2012	2	3	1	Post op check cruciate repair	3	Post op check cruciate repair	3	leg wound 5/10th lame (shampoo)	manage/ther tx
			2	recheck ears		recheck ears		ears fine	nothing
			3	booster /vacc		booster /vacc		vacc	prop tx
15/02/2012	3	2	1	3rd cartophen injection	1	3rd cartophen injection	1	3rd cartophen injection	manage

Date	N	Observation	Order of presentation by client	Brief description	Extraction (Vet 1 assessed)	Brief description	Extraction (Vet 2 assessed)	brief description	Outcome
			2	review of prev prob anal glands found to be normal					nothing
15/02/2012	4	1	1	wound on thorax from bite	1	wound on thorax from bite	1	wound on thorax	ther tx
15/02/2012	5	4	1	vacc booster	4	vacc booster	1	vacc booster***	prop tx
			2	still leaking urine (prostate)		still leaking urine (prostate)			nothing
			3	mass on elbow did look like callus		mass on elbow did look like callus			other
			4	post op check stitches out		post op check stitches out			other
15/02/2012	6	2	1	vacc booster	2	vacc booster	2	vacc booster	prop tx
			2	Conjunctivitis		Conjunctivitis		conjunctivitis	ther tx
15/02/2012	7	2	1	booster vacc	2	booster vacc	2	booster vacc	prop tx
			2	eyelid ward		eyelid wart		eyelid wart	nothing

Date	N	Observation	Order of presentation by client	Brief description	Extraction (vet 1 assessed)	Brief description	Extraction (Vet 2 assessed)	brief description	Outcome
15/02/2012	8	3	1	first vacc/worm	3	first vacc/worm	3	first vacc/worm	prop tx
			2	has fleas		has fleas		has fleas	ther tx
			3	pu/pd		pu/pd		pu/pd	work up
15/02/2012	9	3	1	first vacc/worm	3	first vacc/worm	3	first vacc/worm	prop tx
			2	has fleas		has fleas		has fleas	ther tx
			3	pu/pd		pu/pd		pu/pd	work up
15/02/2012	10	1	1	poss eaten mouse bait	1	poss eaten mouse bait	1	poss eaten mouse bait	ther tx
16/02/2012	1	4	1	booster vacc	1	booster vacc	1	booster vacc	prop tx
			2	had tick on face owner removed - lump left behind					nothing
			3	review of umbilical hernia					0

Date	N	Observation	Order of presentation by client	Brief description	Extraction (Vet 1 assessed)	Brief description	Extraction (Vet 2 assessed)	brief description	Outcome
			4	new concern - thin? Weight found to be normal.					nothing
16/02/2012	2	3	1	booster vacc	2	booster vacc	2	booster vacc	nothing
			2	wound on face		wound on face		wound on face	nothing
			3	new concern - overweight? Watch weight/ diet					manage
16/02/2012	3	4	1	review of polydipsia- recheck after liver probs	1	repeat consult Vomiting + diarhoea stopped. Bloods.	1	recheck V+ and D+ bichem and haem.	manage
			2	review of weight - lost a bit					manage
			3	new problem - trembling/shaking when walking					nothing
			4	new problem - skin lump/wart					nothing

Date	N	Observation	Order of presentation by client	Brief description	Extraction (Vet 1 assessed)	Brief description	Extraction (Vet 2 assessed)	brief description	Outcome
16/02/2012	4	5	1	Bilateral otitis	2	Bilateral otitis	2	Bilateral otitis	manage
			2	vacc booster		vacc booster		vacc booster	prop tx
			3	new concern - loses urine when putting lead on to go for a walk					nothing
			4	New concern - Fussy with food					nothing
			5	new concern - overweight					manage
16/02/2012	5	5	1	vacc/worm	4	vacc/worm	4	vacc/worm	prop tx
			2	Review of arthritis		arthritis		arthritis	Ther. TX
			3	New problem - cataracts		cataracts		cataracts	nothing
			4	review of heart murmur		heart murmur		heart murmur	nothing

Date	N	Observation	Order of presentation by client	Brief description	Extraction (Vet 1 assessed)	Brief description	Extraction (Vet 2 assessed)	brief description	Outcome
			5	review of weight overweight previously - now ok					nothing
16/02/2012	6	3	1	vacc/worm/flea	2	vacc/worm/flea	2	vacc/worm/flea	prop tx
			2	cycst/mass on opcipital crest		cycst/mass on opcipital crest		cycst/mass on opcipital crest	nothing
			3	tick removed by owner					prop tx
16/02/2012	7	3	1	Clip nails	2	Clip nails	1	Clip nails	nothing
			2	awkward with other dogs		awkward with other dogs			manage
			3	New concern - dry skin?					nothing
17/02/2012	1	2	1	Vitamin B injected (4/4)	2	Vitamin B injected (4/4).	2	Vitamin B injected (4/4).	manage
			2	Review of weight/weight loss		Low fat diet		Low fat diet	manage
17/02/2012	2	1	1	post op check	1	post op check	1	post op check	nothing

Date	N	Observation	Order of presentation by client	Brief description	Extraction (Vet 1 assessed)	Brief description	Extraction (Vet 2 assessed)	brief description	Outcome
17/02/2012	3	2	1	rabies blood serology	1	rabies blood serology	1	rabies blood serology	prop tx
			2	tick seen under chin					0
17/02/2012	4	2	1	rabies blood serology	1	rabies blood serology	1	rabies blood serology	work up
			2	Review of lame - still limping nx sx on leg done graft					manage
17/02/2012	5	3	1	worm/vacc	1	worm/vacc	1	worm/vacc	prop tx
			2	Review of prev ingested rat bait					nothing
			3	microchip placement query - no scanner available					nothing
17/02/2012	6	1	1	Euthanasia	1	Euthanasia	1	Euthanasia	euth
18/02/2012	1	5	1	Booster/vaccination/worm	2	Booster/vaccination/worm	2	vaccination/worm	ther tx

Date	N	Observation	Order of presentation by client	Brief description	Extraction (Vet 1 assessed)	Brief description	Extraction (Vet 2 assessed)	brief description	Outcome
			2	Ears OK		Ears OK		ears - NAD	nothing
			3	New concern - dull coat?					manage
			4	New concern - Weight gain					manage
			5	New concern - Grumpy towards other dogs (aggression)					nothing
18/02/2012	2	2	1	Booster/vaccination/worm	2	Booster/vaccination/worm	2	Booster/vaccination/worm	prop tx
			2	Ears OK		Ears OK		Ears - OK	manage
18/02/2012	3	4	1	vacc/worm/flea	2	vacc/worm/flea	2	vacc/worm/flea	prop tx
			2	tartar on teeth		tartar on teeth		tartar on teeth	nothing
			3	behaviour problem - anxious around men and when travelling					nothing
			4	Moulting					nothing

Date	N	Observation	Order of presentation by client	Brief description	Extraction (Vet 1 assessed)	Brief description	Extraction (Vet 2 assessed)	brief description	Outcome
18/02/2012	4	3	1	Review of weightloss, PU/PD downhill, CRF?	2	Review of weightloss, CRF?	2	Review of weightloss	ther tx
			2	Blood pressure check, Blind, Detatched retina		Blood pressure check		Blood pressure check	ther tx
			3	Abnormal bloood result					ther tx
18/02/2012	5	3	1	Conjuntivitis	2	Conjuntivitis	2	Conjuntivitis	ther tx
			2	Enquiry about pet passport		Enquiry about pet passport		Enquiry about pet passport	other
			3	New concern - Weightgain. Normal at present					nothing

Ther tx = Therapeutic Treatment, Prop tx = Prophylactic Treatment, Nothing = No action taken, Manage = Management of an existing problem, euth = Euthanasia. 0 = no data recorded during observation, Other = Proceedure eg. Stitches out. Work up = Test results to be prepared based on analytical results.

Appendix 11 Data for problem analysis by veterinarian

Observation	Extraction	vet
3	2	GS (1)
1	1	GS
3	2	GS
2	2	GS
1	1	GS
5	2	GS
4	1	GS
3	3	GS
2	1	GS
1	1	GS
4	4	GS
2	2	GS
2	2	GS
3	3	GS
3	3	GS
1	1	GS
4	1	GS
3	2	GS
4	1	GS
5	2	GS
5	4	GS
3	2	GS
2	1	GS
1	1	GS
2	1	GS
2	1	GS
3	1	GS
1	1	GS
3	2	GS
3	2	GS
	30	Sum
81	53	Count of problems

Observation	Extraction	vet
3	2	PS (2)
1	1	PS
3	2	PS
5	2	PS
2	2	PS
4	2	PS
	6	Sum
18	11	Count of problems

http://www.rcvs.org.uk/advice-and-guidance/code-of-professional-conduct-for-veterinary-surgeons/

2. Veterinary surgeons and clients

2.1 Veterinary surgeons must be open and honest with clients and respect their needs and requirements.

2.2 Veterinary surgeons must provide independent and impartial advice and inform a client of any conflict of interest.

2.3 Veterinary surgeons must provide appropriate information to clients about the practice, including the costs of services and medicines.

2.4 Veterinary surgeons must communicate effectively with clients and ensure informed consent is obtained before treatments or procedures are carried out.

2.5 Veterinary surgeons must keep clear, accurate and detailed clinical and client records.

2.6 Veterinary surgeons must not disclose information about a client or the client's animals to a third party, unless the client gives permission or animal welfare or the public interest may be compromised.

2.7 Veterinary surgeons must respond promptly, fully and courteously to clients' complaints and criticism.

13. Clinical and client records

Updated 12 April 2012

13.1 Clinical and client records should include details of examination, treatment administered, procedures undertaken, medication prescribed and/or supplied, the results of any diagnostic or laboratory tests (including, for example, radiograph, ultrasound or electrocardiogram images or scans), provisional or confirmed diagnoses, and advice given to the client. It is prudent to include plans for future treatment or investigations, details of proposed follow-up care or advice, notes of telephone conversations, fee estimates or quotations, consents given or withheld and contact details. Ideally, client financial information should be recorded separately from clinical records.

13.2 The utmost care is essential in writing case notes or recording a client's personal details to ensure that they are accurate and that the notes are comprehensible and legible. Clinical and client records should be objective and factual, and veterinary surgeons should avoid making personal observations or assumptions about a client's motivation, financial circumstances or other matters.

13.3 Clinical and client records including radiographic images and similar documents, are the property of, and should be retained by, veterinary surgeons in the interests of animal welfare and for their own protection.

13.4 Copies with a relevant clinical history should be passed on request to a colleague taking over the case.

13.5 Where a client has been specifically charged and has paid for radiographic images or other reports, they are legally entitled to them. A practice may choose to make it clear to clients that they are not charged for radiographs or laboratory reports, but for diagnosis or advice only.

13.6 The Data Protection Act 1998 gives anyone the right to be informed about any personal data relating to themselves on payment of an administration charge. At the request of a client, veterinary surgeons must provide copies of any relevant clinical and client records, including radiographic images and similar documents. This also includes relevant records which have come from other practices, if they relate to the same animal and the same client, but does not include records which relate to the same animal but a different client.

www.rcvs.org.uk

Continuing Professional Development (CPD) for vets

The *RCVS Code of Professional Conduct for Veterinary Surgeons* makes it very clear that veterinary surgeons have a responsibility to ensure that they maintain and develop the knowledge and skills relevant to their professional practice and competence. The *Code* also requires veterinary surgeons to provide the RCVS with their CPD records when requested to do so. CPD is the personal obligation of all responsible veterinary surgeons and should be seen as the continuous progression of capability and competence.

The recommended minimum CPD is 105 hours over a rolling three year period with an average of 35 hours per year. It is appreciated that most veterinary surgeons will do considerably more than this.

What counts as CPD?

Broadly speaking, any activities you undertake in order to further your professional competence as part of a planned development programme can be counted towards your CPD. You do not have to participate in 35 hours of face to face learning delivered by an external CPD provider.

Although this is not an exhaustive list, appropriate activities may include:

- Ad hoc, undocumented private study (up to 10 hours per year)
- Clinical audit activity
- Discussion group - informal learning set
- Distance learning - on-line/ formal (assessed and/or moderated by a third party)
- Distance learning - on-line/informal (not assessed)
- Distance learning - webinars
- Lecture by external provider
- Mentoring or being mentored
- Practical training - clinical skills lab
- Practical training - wet lab
- Preparing a new lecture/presentation
- Project - working on a new project/in a new area of work
- Reading - planned and documented private study/reading
- Research - clinical
- Research - scientific
- Research - veterinary business
- Secondment to another work place
- 'Seeing practice' - work-based observation
- Seminar/workshop - external
- Studying for an external qualification
- Training - in house
- Training as examiner/assessor

Appendix 14 Presentation categories for abstracts published in BSAVA proceedings 2001 – 2010 (n = 111).

1. Abdominal Imaging
2. Abdominal Surgery
3. Anaesthesia
4. Anaesthesia and Analgesia
5. Articular Fractures
6. Articular Surgery
7. Arthropod-Borne Diseases
8. Birds
9. Caged birds
10. Avian Medicine
11. Behaviour
12. Behaviour: Exotics
13. Cardiology
14. Cardiology: Acquired
15. Cardiology: Congenital
16. Clinical Pathology
17. Clinical Pathology: Biochemistry
18. Clinical Pathology: Cytology
19. Critical Care
20. Dentistry
21. Dermatology
22. Dermatology : Atopy
23. ENT Surgery
24. Oral Surgery
25. Nasal disease
26. Ear disease
27. Oropharyngeal surgery
28. Emergency and Critical Care
29. Endocrinology
30. Endocrine
31. Endoscopy
32. Evidence-based medicine
33. Exotics
34. Exotics: Clinical Pathology
35. Exotics: Emergencies
36. Chelonians and Invertebrates
37. Wildlife Casualties
38. Farm pets
39. Fish
40. Small Mammals
41. Feline / Cat
42. Feline Emergencies
43. Feline Infectious Diseases, Recent developments
44. Feline Behaviour
45. Feline Endocrinology
46. Feline Geriatric Medicine
47. Feline Infectious Diseases
48. Feline Behaviour
49. Feline medicine
50. Feline Orthopaedics
51. External Fixation
52. Forensics
53. Gastroenterology
54. GI Surgery
55. Intestinal
56. Geriatrics
57. Haematology
58. Haemostatis
59. Head and Neck Surgery
60. Hepatology
61. Pancreas
62. Imaging
63. Thoracic Imaging
64. Musculoskeletal Imaging
65. Immunology
66. Infection and Vaccination
67. Infectious disease
68. Lameness
69. Nephrology
70. Neurology
71. Neurobiology
72. Neuromuscular Disease
73. Obstetrics & Neonatology
74. Oncology
75. Ophthalmology
76. Vision and Blindness
77. Orthopaedics
78. Orthopaedics: Problems of skeletally mature dogs
79. Orthopaedics: Problems of the immature dog

80. Orthopaedics: Hip Disease
81. Orthopaedics: Conditions of bone
82. Orthopaedics: fractures
83. Orthopaedics (arthrology)
84. Rabbits
85. Rabbits & ferrets
86. Reptiles
87. Reproduction
88. Respiratory
89. Respiratory Medicine
90. Respiratory disease
91. Seizures
92. Soft Tissue: GI Tract
93. Soft Tissue: Refresher
94. Soft Tissue Surgery
95. Sepsis in Surgery
96. Spinal Disease
97. Spinal surgery
98. Surgical Emergencies
99. Tendons, Ligaments, Muscles & the Stifle
100. Theraputics
101. Thoracic Surgery
102. Toxicoses
103. Transfusion medicine
104. Urology
105. Urinary System
106. Urethral Sphincter - Mechanism Incompetence
107. Vascular Disease
108. Welfare
109. Wound Healing and Reconstruction
110. Wound Management
111. Wounds

1. Systematic Reviews and Meta-analysis; A systematic review using a pre-defined method comparing and amalgamating the results of many trials with a similar method to allow drect comparison.
2. Blinded or non-blinded Randomized controlled trial. Animals randomised into groups for either treatment/ intervention verses a control, no treatment / intervention)
3. Clinical trial without randomisation but with a control group (Animals split into two groups for intervention - test & control, balanced for age sex & so on but not randomised)
4. Clinical trial without control but randomised; a group of animals randomised into two groups with more than one treatment or intervention given and outcome compared (eg. a diet with two levels of carbohydrate to look for improvement in condition)
5. Cross-over clinical trials – One panel of animals acting as their own controls with wash out period between treatments or intervention.
6. Observational studies - Cohort studies; two groups, one with exposure of interest one without followed overtime and outcomes observed and compared – No intervention.
7. Observational studies - Cross-sectional survey; A section of the population is examined and diseased and non-diseased individuals identified, data is collected and recorded. Snap shot! No intervention
8. Observational studies - Case-control study. Two balanced groups are identified one with the condition of interest and one without. Data is examined retrospectively (possibly from records) to identify associations or causal factors with disease – No intervention
9. Case series – a number of independent animals all with a disease of interest or healthy animals receiving intervention or preventive medicine. Data is collected and reported.
10. Single Case reports – an independent animal with an unusual disease of interest or healthy animals receiving intervention or preventive medicine. Data is collected and reported.
11. Expert opinion; Editorials, reports, reviews, ideas, summaries.
12. Comparative study; Research from animals which are not the target. Small domestic animal species such as rats, humans, primates.
13. In-vitro; Laboratory studies using samples collected from the target species.
14. Diagnostic methods or method development; no animal data presented.

15. Clinical research trial; one panel of healthy animals no disease present change to usual routine by intervention of some sort e.g. Diet, and any change measured.
16. Unclassified, including discovery research.

Bibliography

ALLEN, T. & KRUGER, J. 2000. Feline lower urinary tract disease In: Hand MS, Thatcher CD, Remillard RL, et al, eds. *Small Animal Clinical Nutrition,* 4th ed.**,** 689-723.

ALTMAN, D. G. & BLAND, J. M. 1994. Statistics Notes: Diagnostic tests 1: sensitivity and specificity. *British Medical Journal,* 308**,** 1552.

ANHOLT, R., STEPHEN, C. & COPES, R. 2012. Strategies for collaboration in the interdisciplinary field of emerging zoonotic diseases. *Zoonoses and Public Health,* 59**,** 229-240.

ANHOLT, R. M. 2013. *Informatics and the Electronic Medical Record for Syndromic Surveillance in Companion Animals: Development, Application and Utility [ethesis Online]. Accessed: 29th June 2013.* PhD, University of Calgary, Canada.

AVMA COMMITTEE ON INFORMATICS 1992. *Survey on Computers & Electronic Information Servi*ces [Online]. Available: www.avma.org [4 Sep 2013].

BAKER, L. H. 2008. Medical records as a communication tool. *Banfield Veterinary Journal,* Jan-Feb**,** 8-10.

BAKKEN, S. 2001. An informatics infrastructure is essential for evidence-based practice. *Journal of the American Medical Informatics Association,* 8**,** 199-201.

BARTLETT, P. C., VAN BUREN, J. W., NETERER, M. & ZHOU, C. 2010. Disease surveillance and referral bias in the veterinary medical database. *Preventive Veterinary Medicine,* 94**,** 264-271.

BATAL, I. & HAUSKRECHT, M. 2010. 'Mining clinical data using minimal predictive rules' C*onference proceedings*. *American Medical Informatics Association.* Washington, DC.

BERNSTEIN, R. M., HOLLINGWORTH, G. R., VINER, G. & LEMELIN, J. Family practice informatics: research issues in computerized medical records. *Conference Proceedings,* Annual Symposium on Computer Application in Medical Care, 1993. American Medical Informatics Association, pg. 93.

BLACK, C., KAYE, J. A. & JICK, H. 2002. Relation of childhood gastrointestinal disorders to autism: nested case-control study using data from the UK General Practice Research Database. *British Medical Journal,* 325**,** 419-421.

BLUMENTHAL, D. & TAVENNER, M. 2010. The "meaningful use" regulation for electronic health records. *The New England Journal of Medicine.,* Volume 363**,** pg. 501-4.

BODEN, L., PARKIN, T. & THRUSFIELD, M. 2010. A History in Data SVEPM 1982 - 2009. *Conference proceedings; Society for Veterinary Epidemiology and Preventive Medicine (SVEPM) Annual Meeting, Nantes, France.*

BONNETT, B. & REID-SMITH, R. 1996. Critical appraisal meets clinical reality: evaluating evidence in the literature using canine hemangiosarcoma as an example. *The Veterinary clinics of North America. Small animal practice,* 26**,** 39.

CASE, J. T., WILCKE, J. & HAHN, A. Status of animal health information standards in the United States. 2000. International Conference of Animal Health Information Specialists.

CHEN, H., FULLER, S. S., FRIEDMAN, C. & HERSH, W. 2005. Knowledge management, data mining, and text mining in medical informatics. *Medical Informatics.* Springer.

COCHRANE, A. L. 1972. *Effectiveness and Efficiency: Random Reflections on Health Services*, Nuffield Provincial Hospitals Trust, London, UK.

COCKCROFT, P. & HOLMES, M. 2003. *Handbook of evidence-based veterinary medicine,* Oxford, Blackwell publishing.

COE, J. B., ADAMS, C. L. & BONNETT, B. N. 2007. A focus group study of veterinarians' and pet owners' perceptions of the monetary aspects of veterinary care. *Journal of the American Veterinary Medical Association,* 231**,** 1510-1518.

CROMBIE, I. K. 2010. *Pocket Guide to Critical Appraisal: A handbook for healthcare professionals*, Published London, BMJ.

DE LUSIGNAN S, J VAN VLYMEN, N HAGUE & DHOUL, N. 2006. Using computers to identify non-compliant people at increased risk of osteoporotic fractures in general practice: a cross-sectional study. *Osteoporosis International,* 17**,** 7.

DE LUSIGNAN S, VALENTIN, T., CHAN, T., HAGUE, N., WOOD, O., VLYMEN, J. V. & DHOUL, N. 2004. Problems with primary care data quality: osteoporosis as an exemplar. *Informatics in Primary care,* 12**,** 10.

DEAR, J. D., SHIRAKI, R., RUBY, A. L. & WESTROPP, J. L. 2011. Feline urate urolithiasis: a retrospective study of 159 cases. *Journal of Feline Medicine and Surgery,* 13**,** 725-732.

DIESEL, G., PFEIFFER, D., CRISPIN, S. & BRODBELT, D. 2010. Risk factors for tail injuries in dogs in Great Britain. *The Veterinary Record*.

DOLIN, R. H. 1997. Outcome analysis: considerations for an electronic health record. *M.D Computing: Computers in Medical Practice,* 14**,** pg. 50-56.

DOWNES, M., CANTY, M. & MORE, S. 2009. Demography of the pet dog and cat population on the island of Ireland and human factors influencing pet ownership. *Preventive Veterinary Medicine,* 92**,** 140 - 149.

DOWNES, M., CLEGG, T., COLLINS, D., MCGRATH, G. & MORE, S. 2011. The spatial distribution of pet dogs and pet cats on the island of Ireland. *BMC Veterinary Research,* 7**,** 28.

DOWNES, M., DEAN, R., STAVISKY, J., ADAMS, V., GRINDLAY, D. & BRENNAN, M. 2013. Methods used to estimate the size of the owned cat and dog population: a systematic review. *BMC Veterinary Research,* 9**,** 121.

ESTBERG, L., CASE, J. T., WALTERS, R. F., CARDIFF, R. D. & GALUPPO, L. D. 1998. Word search performance for diagnoses of equine surgical colics in free-text electronic patient records. *Preventive Veterinary Medicine,* 34**,** 161-174.

EVANS, J., LANE, D. & HENDY, P. 1974. The profile of small animal practice. *Journal of Small Animal Practice,* 15**,** 595-608.

EVERITT, S. 2011. *Clinical decision making in veterinary practice [eThesis Online] Accessed: 9th Sep 2013.* PhD, University of Nottingham.

EVIDENCE-BASED WORKING GROUP 1992. Evidence-based medicine. A new approach to teaching the practice of medicine. *Journal American Medical Association,* 268.

FAUNT, K., LUND, E. & NOVAK, W. 2007. The power of practice: Harnessing patient outcomes for clinical decision making. *Veterinary Clinics of North America-Small Animal Practice,* 37**,** pg. 521.

FERGUSON, E. A., LITTLEWOOD, J. D., CARLOTTI, D. N., GROVER, R. & NUTTALL, T. 2006. Management of canine atopic dermatitis using the plant extract PYM00217: a randomized, double‐blind, placebo‐controlled clinical study. *Veterinary dermatology,* 17**,** 236-243.

FERRANTI, J. M., MUSSER, R. C., KAWAMOTO, K. & HAMMOND, W. 2006. The clinical document architecture and the continuity of care record: a critical analysis. *Journal of the American Medical Informatics Association,* 13**,** 245-252.

FINN, E., FREEMAN, L. M., RUSH, J. E. & LEE, Y. 2010. The Relationship Between Body Weight, Body Condition, and Survival in Cats with Heart Failure. *Journal of Veterinary Internal Medicine,* 24**,** 1369-1374.

FOLEY, R. N., MURRAY, A. M., LI, S. L., HERZOG, C. A., MCBEAN, A. M., EGGERS, P. W. & COLLINS, A. J. 2005. Chronic kidney disease and the risk for cardiovascular disease, renal replacement, and death in the United States medicare population, 1998 to 1999. *Journal of the American Society of Nephrology,* 16**,** 489-495.

FORRESTER, D. S. & ROUDEBUSH, P. 2007. Evidence-Based Management of Feline Lower Urinary Tract Disease. *Veterinary Clinics of North America: Small Animal Practice,* 37**,** 533-558.

GLICKMAN, L. T., GLICKMAN, N. W., MOORE, G. E., GOLDSTEIN, G. S. & LEWIS, H. B. 2009. Evaluation of the risk of endocarditis and other cardiovascular events on the basis of the severity of periodontal disease in dogs. *Journal of the American Veterinary Medical Association,* 234**,** 486-494.

GLICKMAN, L. T., GLICKMAN, N. W., MOORE, G. E., LUND, E. M., LANTZ, G. C. & PRESSLER, B. M. 2011. Association between chronic azotemic kidney disease and the severity of periodontal disease in dogs. *Preventive Veterinary Medicine,* 99**,** 193-200.

GRAY, C. & CRIPPS, P. 2005. 'Typical'veterinary consultation in the United Kingdom and Ireland. *Veterinary record,* 156**,** 381-381.

GREENHALGH, T. 2010. *How to read a paper: The basics of evidence-based medicine*, BMJI Books, Wiley and Blackwell publishing.

HAHN, A., MARTIN, M., SIEGEL, A. & ELLIS, W. 2003. Sending data to a central repository. *Biomedical sciences instrumentation,* 40**,** 475-479.

HAMMERSLEY, V., HIPPISLEY-COX, J., WILSON, A. & PRINGLE, M. 2002. A comparison of research general practices and their patients with other practices — a cross-sectional survey in Trent. *British Journal of General Practice,* 52**,** 6.

HARMAN, J. S., ROST, K. M., HARLE, C. A. & COOK, R. L. 2012. Electronic medical record availability and primary care depression treatment. *Journal of general internal medicine,* 27**,** 962-967.

HASSEY, A., GERRETT, D. & WILSON, A. 2001. A survey of validity and utility of electronic patient records in a general practice. *British Medical Journal,* 322**,** 1401-5.

HÄYRINEN, K., SARANTO, K. & NYKÄNEN, P. 2008. Definition, structure, content, use and impacts of electronic health records: a review of the research literature. *International journal of medical informatics,* 77**,** 291-304.

HERSH, W. R. 2003a. Information Retrieval: A Health and Biomedical Perspective.

HERSH, W. R. 2003b. *Information Retrieval: A Health and Biomedical Perspective*, New York, Springer.

HERSH, W. R., CAMPBELL, E. M. & MALVEAU, S. E. 1997. Assessing the feasibility of large-scale natural language processing in a corpus of ordinary medical records: a lexical analysis. *Conference Proceedings: The American Medical Iinformatics Association (AMIA) Annual Fall Symposium***,** pg. 580.

HIPPISLEY-COX, J., FIELDING, K. & PRINGLE, M. 1998. Are spouses of patients with hypertension at increased risk of having hypertension? A population-based case-control study. *British Journal of General Practice,* 4.

HIPPISLEY-COX, J. & STABLES, D. 2011. QRESEARCH An ethical high quality general practice database for research.: The University of Nottingham and Egton Medical Informations Systems.

HOLMES, M. 2005. Seeking the truth: the evidence-based approach to veterinary medicine. *The Veterinary Record,* 156**,** 528-30.

HOLMES, M. A. 2007. Evaluation of the Evidence. *The Veterinary clinics of North America. Small animal practice,* 37**,** 447-462.

HOLT, T. A., STABLES, D., HIPPISLEY-COX, J., O'HANLON, S. & MAJEED, A. 2008. Identifying undiagnosed diabetes: cross-sectional survey of 3.6 million patients' electronic records. *British Journal of General Practice,* 58**,** 5.

HORNOF, W. J., BRENTSON, P. R., SELF, J. A. & BALLANCE, D. W. 2001. Development of a complete electronic medical record in an academic

institution. *Journal of the American Veterinary Medical Association,* 218**,** 1771-1775.
HRIPCSAK, G., FRIEDMAN, C., ALDERSON, P. O., DUMOUCHEL, W., JOHNSON, S. B. & CLAYTON, P. D. 1995. Unlocking clinical data from narrative reports: a study of natural language processing. *Annals of internal medicine,* 122**,** 681-688.
JOHNSON, L. M., AMES, T. R., JACKO, J. A. & WATSON, L. A. 2011. The Informatics Imperative in Veterinary Medicine: Collaboration across Disciplines. *Journal of veterinary medical education,* 38**,** 5-9.
JOLLIS, J. G., ANCUKIEWICZ, M., DELONG, E. R., PRYOR, D. B., MUHLBAIER, L. H. & MARK, D. B. 1993. Discordance of databases designed for claims payment versus clinical information systems: implications for outcomes research. *Annals of Internal Medicine,* 119**,** 844-850.
KAPATKIN, A. S. 2007. Outcome - Based Medicine and Its Application in Clinical Surgical Practice. *Veterinary Surgery,* 36**,** 515-518.
KASTELIC, J. P. 2006. Critical evaluation of scientific articles and other sources of information: An introduction to evidence-based veterinary medicine. *Theriogenology,* 66**,** 534-542.
KATEHAKIS, D. G., SFAKIANAKIS, S., TSIKNAKIS, M. & ORPHANOUDAKIS, S. C. 2001. An infrastructure for Integrated Electronic Health Record services: the role of XML (Extensible Markup Language). *Journal of Medical Internet Research,* 3.
KEARSLEY-FLEET, L., O'NEILL, D. G., VOLK, H. A., CHURCH, D. B. & BRODBELT, D. C. 2013. Prevalence and risk factors for canine epilepsy of unknown origin in the UK. *Veterinary Record,* 172**,** 338-338.
LAM, K., PARKIN, T., RIGGS, C. & MORGAN, K. 2007a. Use of free text clinical records in identifying syndromes and analysing health data. *Veterinary Record,* 161**,** 547-551.
LAM, K. H., PARKIN, T. D. H., RIGGS, C. M. & MORGAN, K. L. 2007b. Descriptive analysis of retirement of Thoroughbred racehorses due to tendon injuries at the Hong Kong Jockey Club (1992-2004). *Equine Veterinary Journal,* 39**,** 143-148.
LAM, K. K. H., PARKIN, T. D. H., RIGGS, C. M. & MORGAN, K. L. 2007c. Evaluation of detailed training data to identify risk factors for retirement because of tendon injuries in Thoroughbred racehorses. *American Journal of Veterinary Research,* 68**,** 1188-1197.
LANDIS, J. R. & KOCH, G. G. 1977. The measurement of observer agreement for categorical data. *Biometrics,* 33**,** pg. 159-174.
LANGMAN, M. J. S., CHENG, K. K., GILMAN, E. A. & LANCASHIRE, R. J. 2000. Effect of anti-inflammatory drugs on overall risk of common cancer: case-control study in general practice research database. *British Medical Journal,* 320**,** 1642-1646.

LAWRENSON, R., WILLIAMS, T. & FARMER, R. 1999. Clinical information for research; the use of general practice databases. *Journal of Public Health Medicine,* 21**,** 6.

LEAN, I. J., RABIEE, A. R., DUFFIELD, T. F. & DOHOO, I. R. 2009. Invited review, Use of meta-analysis in animal health and reproduction: Methods and applications. *Journal of Dairy Science,* 92**,** pg. 3545-3565.

LIS, Y. & MANN, R. D. 1995. The VAMP research multi-purpose database in the U.K. *Journal of Clinical Epidemiology,* 48**,** 431-443.

LUND, E. 1997. *Development and evaluation of a model for diagnostic surveillance in companion animal practice.* PhD, The University of Minnesota.

LUND, E. 2008a. The epidemiology of atopic dermatitis. *Banfield Journal,* 4**,** 17-22.

LUND, E. 2008b. Risk for feline lower urinary tract disease. *Banfield Journal,* 4**,** 14.

LUND, E. M., ARMSTRONG, P. J., KIRK, C. A. & KLAUSNER, J. S. 2005. Prevalence and risk factors for obesity in adult cats from private US veterinary practices. *International Journal of Applied Research in Veterinary Medicine,* 3**,** 88-96.

LUND, E. M., ARMSTRONG, P. J., KIRK, C. A. & KLAUSNER, J. S. 2006. Prevalence and risk factors for obesity in adult dogs from private US veterinary practices. *International Journal of Applied Research in Veterinary Medicine,* 4**,** 177.

LUND, E. M., ARMSTRONG, P. J., KIRK, C. A., KOLAR, L. M. & KLAUSNER, J. S. 1999. Health status and population characteristics of dogs and cats examined at private veterinary practices in the United States. *Journal of the American Veterinary Medical Association,* 214**,** 1336-1341.

LUND, E. M., KLAUSNER, J. S., ELLIS, L. B. & WHITNEY, J. R. 1998. PetTerms: a system of standardized nomenclature for companion animal practice. *Online Journal of Veterinary Research,* 2.

MALYNICZ, G. 1998. Evidence-based medicine. *The Veterinary record,* 143**,** 619.

MANT, D. 1997. NHS Executive report, R&D in primary care. London: Department of Health,.

MARSHALL, K. 2008a. Electronic record-keeping systems. *Banfield Veterinary Journal,* May/June**,** 8-10.

MARSHALL, K. 2008b. Electronic veterinary medical records. *Banfield Veterinary Journal,* July/Aug**,** 8-12.

MARSHALL, K. 2008c. The role of electronic records in practice. *Banfield Veterinary Journal,* March/April**,** 8-12.

MATEUS, A., BRODBELT, D., BARBER, N. & STÄRK, K. 2011. Antimicrobial usage in dogs and cats in first opinion veterinary practices in the UK. *Journal of Small Animal Practice,* 52**,** pg. 515-521.

MCALLISTER, D. & WILD, S. 2009. Why should clinicians understand epidemiology? *Postgraduate Medical Journal,* 85**,** 313-315.

MCCURDY, H. 2001. The paperless practice. *Journal of the American Veterinary Medical Association,* Vol 218.

MEYSTRE, S. M., SAVOVA, G. K., KIPPER-SCHULER, K. C. & HURDLE, J. F. 2008. Extracting information from textual documents in the electronic health record: a review of recent research. *IMIA Yearbook of Medical Informatics***,** 128-44.

MOORE, G. & LUND, E. 2009. Disease reporting and surveillance: where do companion animal diseases fit in? *Vet Clin North Am Small Anim Pract.,* 39**,** 225-40.

MOORE, G. E., DESANTIS-KERR, A. C., GUPTILL, L. F., GLICKMAN, N. W., LEWIS, H. B. & GLICKMAN, L. T. 2007. Adverse events after vaccine administration in cats: 2,560 cases (2002-2005). *Journal of the American Veterinary Medical Association,* 231**,** 94-100.

MOORE, G. E., GUPTILL, L. F., WARD, M. P., GLICKMAN, N. W., FAUNT, K. K., LEWIS, H. B. & GLICKMAN, L. T. 2005. Adverse events diagnosed within three days of vaccine administration in dogs. *Journal of the American Veterinary Medical Association,* 227**,** 1102-1108.

MUIR, W. 2003. Is evidence-based medicine our only choice? *Equine Veterinary Journal,* 35**,** 337–8.

MURRAY, J. K., BROWNE, W. J., ROBERTS, M. A., WHITMARSH, A. & GRUFFYDD-JONES, T. J. 2010. Number and ownership profiles of cats and dogs in the UK. *Veterinary Record,* 166**,** 163-168.

NEUBERGER, J. 2000. The educated patient: new challenges for the medical profession. *Journal of Internal Medicine,* 247**,** 6-10.

NICHOLSON, A., TATE, A. R., KOELING, R. & CASSELL, J. A. 2011. What does validation of cases in electronic record databases mean? The potential contribution of free text. *Pharmacoepidemiology and Drug Safety,* 20**,** 321-324.

NIELSEN, T. D., DEAN, R. S., ROBINSON, N. J., MASSEY, A. & BRENNAN, M. L. 2014. Survey of the UK veterinary profession: common species and conditions nominated by veterinarians in practice. *The Veterinary Record***,** pg. 324.

NUTTALL, T. & COLE, L. K. 2007. Evidence ‐ based veterinary dermatology: a systematic review of interventions for treatment of Pseudomonas otitis in dogs. *Veterinary Dermatology,* 18**,** 69-77.

NUTTALL, T., MUELLER, R., BENSIGNOR, E., VERDE, M., NOLI, C., SCHMIDT, V. & RÈME, C. 2009. Efficacy of a 0.0584% hydrocortisone aceponate spray in the management of canine atopic dermatitis: a randomised, double blind, placebo ‐ controlled trial. *Veterinary Dermatology,* 20**,** 191-198.

O'NEILL, D., HENDRICKS, A., SUMMERS, J. & BRODBELT, D. 2012. Primary care veterinary usage of systemic glucocorticoids in cats and dogs in three UK practices. *Journal of Small Animal Practice,* 53**,** 217-222.

O'CONNELL, D. & BONVICINI, K. 2007. Addressing disappointment in veterinary practice. *Veterinary Clinics of North America - Small Anim Pract,* 37**,** 135–49.

OLIVRY, T. & MUELLER, R. 2003. Evidence - based veterinary dermatology: a systematic review of the pharmacotherapy of canine atopic dermatitis. *Veterinary Dermatology,* 14**,** 121-146.

PATTERSON, E. S., COOK, R. I. & RENDER, M. L. 2002. Improving patient safety by identifying side effects from introducing bar coding in medication administration. *Journal of the American Medical Informatics Association,* 9**,** 540-553.

PETRIE, A. & SABIN, C. 2009. *Medical statistics at a glance*, Wiley-Blackwell publishing, Oxford.

PETRIE, A. & WATSON, P. 2006. *Statistics for Veterinary and Animal Science*, Blackwell publishing, London.

QUIN, L. R. E. 2012. *The World Wide Web Consortium, XML Schema Standards Accessed Sep 2012: http://www.w3.org/standards/xml/schema.* [Online].

RADFORD, A., TIERNEY, A., COYNE, K., DAWSON, S., NOBLE, P. J. & GASKELL, R. 2010. National surveillance of small animal disease in the UK. *Veterinary Record,* 166**,** 471-472.

RAMSEY, I. 2011. Clinical Research Abstracts. Best kept Congress secret? *Companion.* BSAVA Publishing,.

RICHARD D NEAL, HEYWOOD, P. L. & MORLEY, S. 1996. Real world data—retrieval and validation of consultation data from four general practices. *Family Practice,* 13**,** 7.

ROBERTSON-SMITH, G., ROBINSON, D., HICKS, B., KHAMBHAITA, P. & HAYDAY, S. 2010. *The 2010 RCVS Survey of the UK Veterinary and Veterinary Nursing Professions. Accessed 21st Jun 2013 at www.rcvs.org.uk/document-library/rcvs-survey-of-the-veterinary-professions-2010/* [Online].

ROBINSON, D. & HOOKER, H. 2006. The UK Veterinary Profession in 2006: The Findings of a Survey of the Profession Conducted by the Royal College of Veterinary Surgeons. Royal College of Veterinary Surgeons.

ROBINSON, J., BARRETT, J., EVANS, D. & WILLIAMS, J. 'The state of the nation' *Conference proceedings:* AMIA Annual Symposium. 2008.

RODNICK, J. 1988. Computers and family practice: tools of the trade. *Family Medicine,* 20**,** 16, 68.

RODRIGUEZ, L. A. G. & GUTTHANN, S. P. 1998. Use of the UK General Practice Research Database for pharmacoepidemiology. *British Journal of Clinical Pharmacology,* 45**,** 419-425.

ROTHWELL, D., COTE, R., CORDEAU, J. & BOISVERT, M. 1993. Developing a standard data structure for medical language--the SNOMED proposal. *Conference Proceedings; The Annual Symposium on Computer Application in Medical Care*, pg. 695.

SACKETT, D. L., ROSENBERG, W. M., GRAY, J., HAYNES, R. B. & RICHARDSON, W. S. 1996. Evidence based medicine: what it is and what it isn't. *BMJ: British Medical Journal,* 312**,** 71.

SACKETT DL, R. W., ROSENBERG W, HAYNES RB. 2000. *Evidence-based medicine-how to practice and teach*, 2nd edition. New York: Churchill Livingstone.

SACKETT, D. L. R., W.M. 1995. The need for evidence-based medicine. *Journal of the Royal Society of Medicine,* 88**,** 620–624.

SANTAMARIA, S. L. & ZIMMERMAN, K. L. 2011. Uses of Informatics to Solve Real World Problems in Veterinary Medicine. *Journal of Veterinary Medical Education,* 38**,** 103-109.

SCHERER R, W., LANGENBERG, P. & VON ELM, E. 2007. Full publication of results initially presented in abstracts. *Cochrane Database of Systematic Reviews, Accessed 18th Jun 2013.* [Online]. Available: http://www.mrw.interscience.wiley.com/cochrane/clsysrev/articles/MR000005/frame.html.

SCHMIDT, P. L. 2007. Evidence-based veterinary medicine: Evolution, revolution, or repackaging of veterinary practice? *Veterinary Clinics of North America: Small Animal Practice,* 37**,** 409-417.

SHEARER, P. 2013. Critically Appraised Topic: Diet and Kidney Disease, renal diets may improve quality of life in cats with acute or chronic kidney disease. Accessed 31 Aug 2013. *Banfield Journal* [Online].

SHORTCLIFFE, E. H. & BLOIS, M. S. 2003. *The Computer Meets Medicine and Biology: Energence of a Discipline,* New York, Springer-Verlag.

SINGLETON, B. 1999. History of the BSAVA. Accessed 27 Jun 2013. *http://www.bsava.com/Home/HistoryoftheBSAVA* [Online].

SLATER, M. & BOOTHE, D. 1995. Integrating epidemiologic research and veterinary practice: ideas and experiences. *Conference proceedings: Society for Veterinary Epidemiology and Preventive Medicine, UK.*

SLATER, M. R. 2010. Epidemiological research and evidence based medicine: How do they fit and for whom. *Preventive Veterinary Medicine,* 97**,** 157-164.

SMITH, C. J. P., GRIBBIN, J., CHALLEN, K. B. & HUBBARD, R. B. 2008. The impact of the 2004 NICE guideline and 2003 General Medical Services contract on COPD in primary care in the UK. *QJM: An International Journal of Medicine,* 101**,** 145-153.

SNEDEKER, K. G., CAMPBELL, M., TOTTON, S. C., GUTHRIE, A. & SARGEANT, J. M. 2010. Comparison of outcomes and other variables between conference abstracts and subsequent peer-reviewed papers involving pre-harvest or abattoir-level interventions against foodborne pathogens. *Preventive Veterinary Medicine,* 97**,** 67-76.

SPYNS, P. 1996. Natural Language Processing. *Methods of information in medicine,* 35**,** 285-301.

STAVISKY, J., BRENNAN, M., DOWNES, M. & DEAN, R. 2012. Demographics and economic burden of un-owned cats and dogs in the UK: results of a 2010 census. *BMC Veterinary Research,* 8**,** 163.

STEIN, H. D., NADKARNI, P., ERDOS, J. & MILLER, P. L. 2000. Exploring the degree of concordance of coded and textual data in answering clinical queries from a clinical data repository. *Journal of the American Medical Informatics Association,* 7**,** 42-54.

STOTT, N. C. & DAVIS, R. H. 1979. The exceptional potential in each primary care consultation. *Journal of the Royal College of General Practitioners,* 29**,** 201-5.

SUMMERS, J., HENDRICKS, A., LOEFFLER, A., STÄRK, K. & BRODBELT, D. 2009. Small animal diseases seen in veterinary practice in the UK. *The Veterinary Record, August 8*.

SUMMERS, J. F., HENDRICKS, A., LOEFFLER, A., STÄRK, K. & BRODBELT, D. C. 2010. The Efficacy of Antimicrobials for the Treatment of Canine Pyoderma in the UK: First opinion practice pilot study findings. *Conference proceedings: Society for Veterinary Epidemiology and Preventive Medicine, Nantes, France.*

SUSAN S. JICK, D. S., JAMES A. KAYE, M.D., DR.P.H., CATHERINE VASILAKIS-SCARAMOZZA, M.P.H.,, LUIS A. GARCÍA RODRÍGUEZ, M. D., M.SC., ANA RUIGÓMEZ, M.D., PH.D., CHRISTOPH R. MEIER, PH.D., M.SC., & RAYMOND G. SCHLIENGER, P. D., CORRI BLACK, M.B., AND HERSHEL JICK, M.D. 2003. Validity of the General Practice Research Database. *Pharmacotherapy* 23.

SZOLOVITS, P. 2003. Adding a medical lexicon to an English Parser. *Conference proceedings: AMIA Annual Symposium***,** 639-43.

TATE, A. R., MARTIN, A. G. R., ALI, A. & CASSELL, J. A. 2011. Using free text information to explore how and when GPs code a diagnosis of ovarian cancer: an observational study using primary care records of patients with ovarian cancer. *Bmj Open,* 1.

THIRU, K., DE LUSIGNAN, S. & HAGUE, N. 1999. Have the completeness and accuracy of computer medical records in general practice improved in the last five years? The report of a two practice pilot study. *Health Informatics Journal,* 5.

THRUSFIELD, M. 2013. *Veterinary epidemiology third edition*, Blackwell science. Oxford.

TIM A HOLT, DAVID STABLES, JULIA HIPPISLEY-COX, O'HANLON, S. & MAJEED, A. 2008. Identifying undiagnosed diabetes: cross-sectional survey of 3.6 million patients' electronic records. *British Journal of General Practice,* 58**,** 5.

TUNIS, S. R., STRYER, D. B. & CLANCY, C. M. 2003. Practical clinical trials - Increasing the value of clinical research for decision making in clinical and health policy. *Journal of the American Medical Association,* 290**,** 1624-1632.

UNKNOWN. 1989. Future directions for veterinary medicine: The Pew report Pew National Veterinary Education Program. *Canadian Veterinary Journal,* 30**,** 472-6.

UZUNER, O., SZOLOVITS, P. & KOHANE, I. 2006. i2b2 workshop on natural language processing challenges for clinical records. *Conference proceedings: The Fall Symposium of the American Medical Informatics Association*.

VAN BEMMEL, J. H. & MUSEN, M. A. 1997. *The Patient record,* Bohn Stafleu Van Loghum, Springer.

VANDEWEERD, J.-M., VANDEWEERD, S., GUSTIN, C., KEESEMAECKER, G., CAMBIER, C., CLEGG, P., SAEGERMAN, C., REDA, A., PERRENOUD, P. & GUSTIN, P. 2012a. Understanding Veterinary Practitioners' Decision-Making Process: Implications for Veterinary Medical Education. *Journal of Veterinary Medical Education,* 39**,** 142-151.

VANDEWEERD, J. M., CLEGG, P., BUCZINSKI, S. & WOMACK, C. 2012b. How Can Veterinarians Base Their Medical Decisions on the Best Available Scientific Evidence? *Veterinary Clinics of North America - Food Animal Practice*.

VANDEWEERD, J. M., KIRSCHVINK, N., CLEGG, P., VANDENPUT, S., GUSTIN, P. & SAEGERMAN, C. 2012c. Is evidence-based medicine so evident in veterinary research and practice? History, obstacles and perspectives. *Veterinary Journal,* 191**,** 28-34.

VENOM CODING GROUP. *Clinical coding systems for veterinary practice. Accessed 2 Jul 2013* [Online]. http://www.venomcoding.org/VeNom/Welcome.html.

VMDB. *The Veterinary Medical Database (VMDB). Accessed 18 Jul 2013* [Online]. Available: http://www.vmdb.org/.

WALLEY, T. & MANTGANI, A. 1997. The UK General Practice Research Database. *The Lancet,* 350**,** 1097-1099.

WANG, Z., SHAH, A. D., TATE, A. R., DENAXAS, S., SHAWE-TAYLOR, J. & HEMINGWAY, H. 2012. Extracting Diagnoses and Investigation Results from Unstructured Text in Electronic Health Records by Semi-Supervised Machine Learning. *Plos One,* 7.

WHITELAW, F. G., TAYLOR, R. J., NEVIN, S. L., TAYLOR, M. W., MILNE, R. M. & WATT, A. H. 1996. Completeness and accuracy of morbidity and repeat prescribing records held on general practice computers in Scotland. *British Journal of General Practice,* 46.

WILSON, E. B. 1927. Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association,* 22**,** 209-212.

WISE, J. C. & YANG, J. J. 1992. Computer use and applications in veterinary medicine. *Journal of the American Veterinary Medical Association,* 201**,** 1352-1354.

ZELINGHER, J., RIND, D. M., CARABALLO, E., TUTTLE, M. S., OLSON, N. E. & SAFRAN, C. 1995. Categorization of free-text problem lists: an effective method of capturing clinical data. *Conference proceedings: The 19th Annual Symposium on Computer Applications in Medical Care***,** 416–420.