Accuracy of the electronic patient record in first opinion veterinary practice

Julie Jones-Diette[1*], Natalie Robinson[1], Malcolm Cobb[2], Marnie Brennan[1], Rachel S Dean[1]

[1] Centre for Evidence-based Veterinary Science, School of Veterinary Medicine & Science, University of Nottingham, Nottingham, UK. LE 12 5RD

[2] Deputy Head of School, School of Veterinary Medicine & Science, University of Nottingham, Nottingham, UK. LE12 5RD

Corresponding author: Dr Rachel Dean

Email:Rachel.Dean@nottingham.ac.uk

Alternative email: rdeancevm@gmail.com

Address: Centre for Evidence-based Veterinary Science, School of Veterinary Medicine & Science, University of Nottingham, Nottingham, UK. LE12 5RD.

Telephone: +44 115 951 6575

* Current address, Centre for Reviews and Dissemination, University of York, York UK

Abstract

The use of electronic patient records (EPRs) in veterinary research is becoming more common place. To date no-one has investigated how accurately and completely they represent the clinical interactions that happen between veterinary professionals, and their clients and patients. The aim of this study was to compare data extracted from consultations within EPRs with data gathered by direct observation of the same consultation. A secondary aim was to establish the inter-rater reliability of two researchers who examined the data extracted from the EPRs. A convenience sample of 36 small animal consultations undertaken by 2 veterinary surgeons at a mixed veterinary practice in the United Kingdom was studied. All 36 consultations were observed by a single researcher using a standardised data collection tool. The information recorded in the EPRs was extracted from the Practice Management Software (PMS) systems using a validated XML schema. The XML extracted data was then converted into the same format as the observed data by two independent researchers who examined the extracted information and recorded their findings using the same tool as for the observation. The issues discussed and any action taken relating to those problems recorded in the observed and extracted datasets were then compared. In addition the inter-rater reliability of the two researchers who examined the extracted data was assessed.

Only 64.4% of the observed problems discussed during the consultations were recorded in the EPR. The type of problem, who raised the problem and at what point in the consultation the problem was raised significantly affected whether the problem was recorded or not in the EPR. Only 58.3% of observed actions taken during the consultations were recorded in the EPR and the type of action significantly affected whether it would be recorded or not. There was moderate agreement between the two researchers who examined the extracted data. This is the first study that examines how much of the activity that occurs in small animal consultations is recorded in the EPR. Understanding the completeness, reliability and validity of EPRs is vital if they are to continue to be used for clinical research and the results to direct clinical care.

Key words: electronic patient record, consultation, validity, reliability

1. INTRODUCTION

First opinion small animal veterinary practices are a valuable source of research data and, different methods have been used to harness this data (Lund, 2015). Some studies have used questionnaires (Robotham and Green, 2004; Buchan et al., 2011; Nielsen et al., 2014), while others have involved direct observation of consultations (Hill et al., 2006; Robinson et al., 2015a). Other have collected the clinical information digitally recorded within the practice known as the electronic patient records (EPRs) stored within the Practice Management Software (PMS) systems (computer software used in practice) often in combination with forced entry fields (Faunt et al., 2007)(), clinical coding (Lund et al., 1999; O'Neill, 2013) or embedded questionnaires (Jones et al., 2014). The direct extraction of EPRs from practice computers could provide a more efficient method of collecting large volumes of data for research. However, PMS systems were designed primarily for accountancy purposes to assist with billing and stock control (Musen and van Bemmel, 1997; Shortcliffe and Blois, 2003) and to support the clinical care of patients, therefore limiting their use for research. Despite this, associated and computer-based technologies such as Natural Language Processing (NLP), Information extraction (IE), Content analysis (CA) and Text mining (TM) can support data extraction, management and analysis (Vandeweerd et al., 2012a; Vandeweerd et al., 2012b). However the suitability of information captured within the EPR may 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). It is therefore important to understand how accurately an EPR reflects the actual consultation, as then can the limitations of using EPRS for research can be fully understood. It has been suggested that patient signalment data recorded in EPRs are relatively consistent (Robinson et al., 2015b), however other studies have found discrepancies occurring relatively frequently (Dean et al., 2013). Recent research has highlighted the complexity of first opinion consultations (Everitt et al., 2013; Robinson et al., 2015a) and it is currently unclear how much of this complexity is recorded in EPRs... In human healthcare it has been shown that not all problems discussed during primary care consultations are recorded (Flocke et al., 2001; Beasley et al., 2004) but to the authors knowledge this has not been

done in veterinary medicine. The aim of this study was to examine the information captured in the PMS system of a first opinion veterinary practice during one week of small animal consultations and to compare to data collected by direct observation during the same consultations. A secondary aim was to compare the inter-rater reliability between researchers when interpreting data extracted from EPRs.

2. MATERIALS AND METHODS

2.1. DATA COLLECTION

A first opinion mixed, but primarily small animal, practice with three veterinary surgeons was selected based on their involvement with previous work with the Centre for Evidence-based Veterinary Medicine (CEVM) (Robinson et al., 2015a). Data was collected during one working week (5 days, 14-18th May 2012) by direct observation of the consultations and by automated EPR extraction using an XML schema (Jones-Diette et al., 2016)

2.1.1 Observational data collection

A data collection tool was developed to allow real-time collection of data by a veterinary researcher experienced in observing veterinary consultations (author NR) (Robinson et al., 2015a). The date and time of consultation, patient signalment and consulting veterinary surgeon ID were all recorded. For each consultation, type of consultation was recorded as being either a preventive medicine consultation or a specific health problem consultation (Robinson et al., 2016). Data were gathered on all problems discussed during each consultation, in the order they were discussed. A problem was defined as 'any two-way discussion between owner/carer and veterinary surgeon regarding any aspect of the patients' health and wellbeing'. Each problem was categorised as having been raised by the veterinary surgeon, by the owner or by a prompt, for example a vaccination or worming reminder card. For each problem recorded, the type of problem was categorised as being a new health problem, a pre-existing health problem or a preventive medicine problem. In addition for each problem, the action or actions taken for the problem was also categorised. The action categories were: therapeutic treatment; prophylactic treatment; management; work-up; euthanasia; no action; other.

96 The tool was initially developed in Microsoft® Office Word 2010 then transferred to Cardiff Teleform®
97 Version 10.5.1 (Verity Inc., Cambridge). Completed forms were then scanned, verified and exported
98 to a Microsoft® Office Access 2010 database for analysis. Development, piloting, reliability and
99 utilisation of the tool have been reported in more detail previously (Robinson et al., 2015a, 2016).
100 2.1.2 *Extracted XML data collection*
101 An XML schema (Jones-Diette et al., 2016), was used to extract the information entered into the EPRs
102 during the data collection period. For each entry, the schema extracted information about the practice
103 (ID number only), the patient (ID and signalment) and the consultation or transaction details (time,
104 date, person entering the data e.g. consulting veterinary surgeon, free text including clinical notes,
105 diagnosis codes where available and treatments prescribed). All entries recorded during the data
106 collection week were extracted by the practice and forwarded via email as an XML data file for
107 analysis.
108 *2.1.3 Data comparison*
109 The consultations that were directly observed were then matched to the EPRs collected via XML
110 extraction. This was done by matching the date and time of the consultation, patient signalment data
111 (e.g. breed, age, sex and neuter status) and the ID of the consulting veterinary surgeon. To enable
112 direct comparison of the extracted data with the observed data, the extracted information was
113 entered into the same direct observer data collection tool by the same researcher (NR). The observed
114 dataset was then compared with the extracted information, with the observed dataset used to
115 establish what was and wasn't recorded in the EPR (comparison 1). Of particular interest was whether
116 each problem observed was recorded in the EPR, and also whether each action observed was recorded
117 in the EPR. For problems which were recorded in the EPR and had an action of 'No action' (as
118 ascertained by observation), an action of 'No action' was considered to also be recorded in the EPR
119 provided it was clear that no action had been taken or was necessary (for example statements such
120 as 'monitor for now' or 'all fine now'). Then XML extracted data was also entered in to the data

collection tool by a second researcher (RD) to establish inter-rater reliability of interpreting data extracted from an EPR (comparison 2).

2.2 Data Analysis

2.2.1 Comparison 1: Direct comparison of information captured by observation and from the extracted data by a single researcher (NR)

By using the anonymous veterinary surgeon code extracted from the PMS system, it was possible to calculate the number of problems recorded by the veterinary surgeons in the extracted data and compare this to the number of problems recorded during observation. It was also possible to determine whether certain variables may have influenced the number of problems recorded in the EPRs (see below). Descriptive statistics were generated using IBM® SPSS® 22. As these are related-samples Wilcoxon signed rank test was conducted to compare the number of problems recorded by direct observation with the number recorded from the XML extracted data for each visit. Chi-square analyses were carried out to compare categorical variables, with comparisons made between whether the problem was recorded in the EPR versus: consulting veterinary surgeon; type of consultation; type of problem; who raised the problem. Chi-square analysis was also carried out to compare whether the action taken was recorded in the EPR versus the type of action taken. Only the four most frequently observed types of action were included in the analysis. A Mantel-Haenszel test of trend was carried out to explore the relationship between whether the problem was recorded in the EPR with the order in which the problem was discussed (for example first problem discussed during the consultation, second problem discussed, and so on, as ascertained by direct observation).Statistical significance was initially set at 0.05, with a Bonferroni correction applied to account for multiple comparisons (Petrie and Sabin, 2009).

2.2.2 Comparison 2: Comparison and inter-rater reliability of examining extracted data by two independent researchers (NR and RD)

Agreement between the two datasets created independently by 2 researchers examining the same extracted data was measured by comparing the number of problems identified by the researchers in

147 the extracted data. The level of agreement was assessed using a linear-weighted Kappa analysis to
148 determine the inter-rater reliability (www.vassarstats.net/kappa.html). A kappa score between 0.60
149 – 0.79 suggests a moderate level of agreement, 0.80 – 0.90 a strong level of agreement and above
150 0.90 an almost perfect level of agreement (McHugh, 2012).

151 **3. RESULTS**

152 During the data collection period, data were entered into the EPRs of 158 animals. In total, 86 of these
153 entries did not involve a small animal consultation, leaving 72 entries recorded by 2 veterinary
154 surgeons (Figure 1). Of the 72 small animal consultations for which data were entered into the PMS,
155 36 consultations with 2 veterinary surgeons had been directly observed (Figure 1).

156 **Figure 1. Flow diagram presenting the number of consultation record extracted from the practice**
157 **computer using the XML schema. The extracted data (n = 158 entries) were then cross-matched with**
158 **data collected by direct observation (n=36 consultation) to allow examination of the data recorded**
159 **into the animal record during the observed consultations**.

160 *3.1 Comparison 1:* Direct comparison of information captured by observation and from the extracted
161 data by a single researcher (NR)

162 The Bonferroni correction resulted in a new significance level of p=0.007. For these 36 animal visits,
163 99 problems were recorded by direct observation with a mean of 2.75 per animal visit (range 1-5and
164 64 problems were recorded in the EPRs for the 36 animal visits, with a mean of 1.78 (range 1-4)
165 problems recorded per visit. Therefore the data recorded in the EPRs was found to represent
166 approximately 64.6% of all problems recorded during direct observation. Significantly more problems
167 per visit were identified by direct observation that were recorded in the clinical notes as assessed
168 using the XML schema ($p<0.001$). If only one problem was discussed this was recorded, while for
169 consultations with more than one problem discussed, all problems discussed were not always
170 recorded (Figure 2).

171 **Figure 2. The number of problems observed and recorded during animal consultations using a**
172 **direct observation method (n = 99 problems) observed by an independent researcher (NR)**

compared with those problems recorded by the consulting veterinarian and extracted from the practice computer using the XML schema as assessed by the same independent researcher (N=64) for each of the 36 consultations.

In total, 30 consultations conducted by Veterinary Surgeon-1 and 6 consultations conducted by Veterinary Surgeon-2 were observed. Veterinary Surgeon-1, recorded 65.4% of problems discussed in the EPR (n=53/81) Veterinary Surgeon-2 recorded 61.1% of problems discussed (n=11/18). The recording behaviour between the two consulting veterinary surgeons was not significantly different (p=0.729). Whether the problem was discussed during a preventive medicine consultation (n=41/64; 64.1% problems recorded) or a specific health problem consultation (n=23/35; 65.7% problems recorded) did not significantly affect whether the problem was recorded (p=0.869).

All problems discussed first during the consultation (n=36/36) were recorded, while all problems discussed fifth (n=5/5) were not recorded in the EPR. Problems discussed earlier in the consultation were significantly more likely to be recorded than problems discussed later in the consultation ($p<0.001$). Recording of the problem in the EPR varied significantly with who had raised the problem ($p<0.001$). Problems prompted by a prompt were most frequently recorded (n=21/21; 100.0% of problems recorded), followed by problems raised by the veterinary surgeon (n=22/30; 73.3%) and then by the owner (n=20/47; 42.6%). Recording of the problem in the EPR also varied significantly with type of problem discussed (p=0.001). Preventive medicine problems were recorded most frequently (n=25/28; 89.3% of problems recorded), followed by pre-existing problems (n=21/32; 65.6%) then new problems (n=18/39; 46.2%).

In total, 103 actions were taken for the 99 problems discussed, with 60/103 (58.3%) of these actions recorded in the EPR and 43/103 (41.7%) not recorded. For 8/64 (12.5%) problems which were recorded in the EPR, the action taken for that problem was not recorded. Recording of the action in the EPR varied significantly with the type of action taken ($p<0.001$). Problems resulting in 'Therapeutic treatment' or 'Prophylactic treatment' were more frequently recorded in the EPR than problems resulting in an action of 'No action' or 'Manage' (Table 1).

Table 1. Differences (p<0.001) in actions recorded in electronic patient records (EPR) versus those not recorded in the EPR by action type for 99 problems recorded by direct observation of 36 small-animal veterinary consultations over a one-week period at a single first-opinion veterinary practice in the UK in 2012 (using data for the four most common action types: therapeutic treatment, prophylactic treatment, no action, and manage.

3.2 Comparison 2: Comparison and inter-rater reliability of examining extracted data by two independent researchers (NR and RD)

The first (NR) and second (RD) reviewers identified 64 (mean 1.78, range 1-4 per visit) and 56 (mean 1.56, range 1-4 per visit) problems from the extracted EPR data across the 36 consultations (Figure 3). The linear-weighted Kappa was found to be K = 0.727 (95% Confidence Interval 0.502-0.952). This means that the level of agreement between the two reviewers for interpretation and subsequent transfer of the extracted data from the EPR to the paper form was considered to be moderate (McHugh, 2012).

Figure 3. Plot to demonstrate the agreement between two independent reviewers (NR & RD) when comparing the number of problems identified by the researchers (NR = 64, RD = 56) in the XML extracted dataset. The level of agreement was assessed using a linear-weighted Kappa analysis to determine the inter-rater reliability, K = 0.727 (95% Confidence Interval 0.502-0.952).

4. DISCUSSION

In this study approximately two thirds of everything discussed in the observed consultations were recorded in the electronic patient records. Factors that seemed to affect what was, and was not, recorded included: the type of problem, when in the consultation the problem was raised, who raised the problem, the total number of problems and whether any action was taken to deal with the problem. This would suggest the complexity of small animal consultations is not currently captured in EPRs. In addition different researchers may interpret EPRs differently. All of this must be considered when using EPRs for research as it affects the reliability of the data.

224 Veterinary consultations are complex and time is limited for decision making completing an EPR
225 (Robinson et al., 2014, Robinson et al., 2015a). Therefore it is not surprising that the EPRs do not
226 contain all of what happens in a consultation. Only two veterinary surgeons were included in this study
227 but they both recorded similar proportions of the observed activity in the consultation. It is possible
228 that time may be one of the factors that affects how much of the consultation is recorded as the
229 consultations which had fewer problems were more completely recorded. Alternatively this could
230 represent a less complex consultation which is easier to recall and record at the end of the
231 consultation. If the ability to recall what is discussed affects what is recorded, it may be expected that
232 problems discussed later in the consultation would be more likely to be recorded than those at the
233 beginning. However in this study the reverse was true, which suggests the recording behaviours of
234 veterinary practitioners may be more complicated than just a matter of what they can remember.
235 The person raising the problem in the consultation affected the likelihood of it being recorded as did
236 the type of problem. It is impossible to look at these factors independently as inherently a specific
237 health problem will rarely be initiated by a prompt in the same way a routine preventive medicine
238 problem is (e.g. vaccination). Another factor that affected whether a problem was likely to be recorded
239 in the EPR is what action was taken, with problems requiring therapeutic or prophylactic treatments
240 being more frequently recorded than if no action or management changes were recommended. This
241 may be related to treatments more likely being charged for so this information will be entered into
242 the EPR to allow it to be invoiced. Alternatively it may be that prophylactic and therapeutic treatments
243 are perceived to be more important by the consulting veterinary surgeons than management or using
244 the 'test of time' (waiting to see if the animal gets better with no treatment) options.
245 The inter-rater reliability between the two observers who examined the data extracted from the EPR
246 was found to be moderate. This choice of tool used to examine the data extracted from the EPR may
247 account for some of the variation as one of the researchers (NR) was more familiar with the tool than
248 the other (RD) and this may be why the results differed. Alternatively the way in which an EPR is
249 interpreted may depend on factors such as clinical experience and familiarity with the subject. The

variation in interpretation between individuals needs to be considered when using EPRs for research as bias may be introduced that affects how the findings are analysed and reported. It would therefore be preferable to ensure that at least some of the extracted data used in any research is reviewed by at least two researchers.

There are a number of limitations to this work. A small number of consultations with only two veterinary surgeons were observed and recorded in this study; a much larger sample size from a number of randomly selected practices would have been desirable. The small sample size means the interpretation of the inferential statistics needs to be cautiously done. The observation method (Robinson et al., 2015a), is time consuming and expensive to undertake therefore it is difficult to undertake on a large scale. It also affects the running practice so placing an observer in the practice for a long period of time is not appropriate. It is unclear how representative the sample of consultations used in this study would be of all consultations by all clinicians. In addition a close working relationship with a practice is required to undertake this type of research as a high level of trust is required to not only access the EPRs but to also be present in and observe consultations., hence this practice was not randomly selected. Qualitative work involving all veterinary professionals that contribute to the content of EPRs needs to be undertaken to understand why certain things are or are not recorded and vital if researchers wish to influence what is put in EPRs to increase their validity as a data source for veterinary research. Despite these limitations this research has raised some important issues that warrant further research if the validity of veterinary EPRs for research is to be understood further and the reliance on EPRs for clinical research remains.

In conclusion, this novel study has shown that not all of the activity of a consultation is recorded in veterinary EPRs and different researchers may interpret the information obtained from EPRs differently. It is vital that more work looking at the validity and reliability of EPRs is undertaken as using this type of data for clinical research may affect behaviours of clinicians or clinical outcomes for patients.

ACKNOWLEDGEMENTS

276 The Centre for Evidence-based Veterinary Medicine is funded by an unrestricted grant from Elanco
277 Animal Health and the University of Nottingham

278 **REFERENCES**

279 Beasley, J.W., Hankey, T.H., Erickson, R., Stange, K.C., Mundt, M., Elliott, M., Wiesen, P., Bobula, J.,
280 2004. How Many Problems Do Family Physicians Manage at Each Encounter? A WReN Study. Annals
281 of Family Medicine 2, 405-410.
282 Bernstein, R.M., Hollingworth, G.R., Viner, G., Lemelin, J., 1993. Family practice informatics: research
283 issues in computerized medical records. Proceedings of the Annual Symposium on Computer
284 Application in Medical Care, 93-97.
285 Buchan, I., Radford, A., Tierney, A., Coyne, K., Gaskell, R., Noble, P., Dawson, S., Setzkorn, C., Jones,
286 P., Newton, J., Bryan, J., 2011. Small Animal Veterinary Surveillance Network (SAVSNET): a 3-
287 month study with vets in practice., BSAVA 54th Annual Congress Birmingham, UK
288 Dean, R.S., Pfeiffer, D.U., Adams, V.J., 2013. The incidence of feline injection site sarcomas in the
289 United Kingdom. BMC Veterinary Research 9, 1-6.
290 Everitt, S., Pilnick, A., Waring, J., Cobb, M., 2013. The structure of the small animal consultation.
291 Journal of Small Animal Practice 54, 453-458.
292 Faunt, K., Lund, E., Novak, W., 2007. The Power of Practice: Harnessing Patient Outcomes for Clinical
293 Decision Making. Veterinary Clinics of North America: Small Animal Practice 37, 521-532.
294 Flocke, S.A., Frank, S.H., Wenger, D.A., 2001. Addressing multiple problems in the family practice
295 office visit. The Journal of family practice 50, 211-216.
296 Hill, P.B., Lo, A., Eden, C.A.N., Huntley, S., Morey, V., Ramsey, S., Richardson, C., Smith, D.J., Sutton,
297 C., Taylor, M.D., Thorpe, E., Tidmarsh, R., Williams, V., 2006. Survey of the prevalence,
298 diagnosis and treatment of dermatological conditions in small animals in general practice.
299 Veterinary Record 158, 533-539.
300 Jones-Diette, J., Brennan, M.L., Cobb, M., Doit, H., Dean, R.S., 2016. A method for extracting patient
301 record data from practice management software systems used in veterinary practice. *In*
302 *press*.

303 Jones, P.H., Dawson, S., Gaskell, R.M., Coyne, K.P., Tierney, Á., Setzkorn, C., Radford, A.D., Noble,
304 P.J.M., 2014. Surveillance of diarrhoea in small animal practice through the Small Animal
305 Veterinary Surveillance Network (SAVSNET). The Veterinary Journal 201, 412-418.
306 Lund, E.M., 2015. Power of practice: using clinical data to advance veterinary medicine. Veterinary
307 Record 176, 46-47.
308 Lund, E.M., Armstrong, P.J., Kirk, C.A., Kolar, L.M., Klausner, J.S., 1999. Health status and population
309 characteristics of dogs and cats examined at private veterinary practices in the United
310 States. J Am Vet Med Assoc 214, 1336-1341.
311 Musen, M.A., van Bemmel, J.H., 1997. Handbook of medical informatics. Bohn Stafleu Van Loghum
312 Houten.
313 Nielsen, T.D., Dean, R.S., Robinson, N.J., Massey, A., Brennan, M.L., 2014. Survey of the UK
314 veterinary profession: common species and conditions nominated by veterinarians in
315 practice. Veterinary Record 174, 324.
316 O'Neill, D., 2013. Pointing the way to improved welfare for companion animals. Veterinary Record
317 173, 240-242.
318 Robinson, N.J., Dean, R.S., Cobb, M., Brennan, M.L., 2014. Consultation length in first opinion small
319 animal practice. Veterinary Record 175, 486.
320 Robinson, N.J., Brennan, M.L., Cobb, M., Dean, R.S., 2015a. Capturing the complexity of first opinion
321 small animal consultations using direct observation. Veterinary Record, 176, 48.
322 Robinson, N.J., Dean, R.S., Cobb, M., Brennan, M.L., 2015b. Investigating common clinical
323 presentations in first opinion small animal consultations using direct observation. Veterinary
324 Record 176, 463.
325 Robinson, N.J., Brennan, M.L., Cobb, M., Dean, R.S., 2016. Investigating preventive-medicine
326 consultations in first-opinion small-animal practice in the United Kingdom using direct
327 observation. Preventive Veterinary Medicine 124, 69-77.

328 Robotham, J., Green, L.E., 2004. Pilot study to investigate the feasibility of surveillance of small
329 animals in the UK. The Journal of small animal practice 45, 213-218.
330 Shortcliffe, E.H., Blois, M.S., 2003. The Computer Meets Medicine and Biology: Energence of a
331 Discipline. Springer-Verlag New York.
332 Vandeweerd, J.M., Clegg, P., Buczinski, S., 2012a. How can veterinarians base their medical decisions
333 on the best available scientific evidence? The Veterinary clinics of North America. Food
334 animal practice 28, 1-11, vii.
335 Vandeweerd, J.M., Vandeweerd, S., Gustin, C., Keesemaecker, G., Cambier, C., Clegg, P., Saegerman,
336 C., Reda, A., Perrenoud, P., Gustin, P., 2012b. Understanding veterinary practitioners'
337 decision-making process: implications for veterinary medical education. Journal of veterinary
338 medical education 39, 142-151.
339
340

Table 1. Differences ($p<0.001$) in actions recorded in electronic patient records (EPR) versus those not recorded in the EPR by action type for 99 problems recorded by direct observation of 36 small-animal veterinary consultations over a one-week period at a single first-opinion veterinary practice in the UK in 2012 (using data for the four most common action types: therapeutic treatment, prophylactic treatment, no action, and manage.

	Recorded		Not recorded		
Action type	**n**	**%**	**n**	**%**	**Total n**
Therapeutic treatment	15	71.4	6	28.6	21
Prophylactic treatment	22	100.0	0	0.0	22
Manage	6	35.3	11	64.7	17
Work up	4	100.0	0	0.0	3
Other	3	100.0	0	0.0	4
Euthanasia	1	100.0	0	0.0	1
No action	9	25.7	26	74.3	35

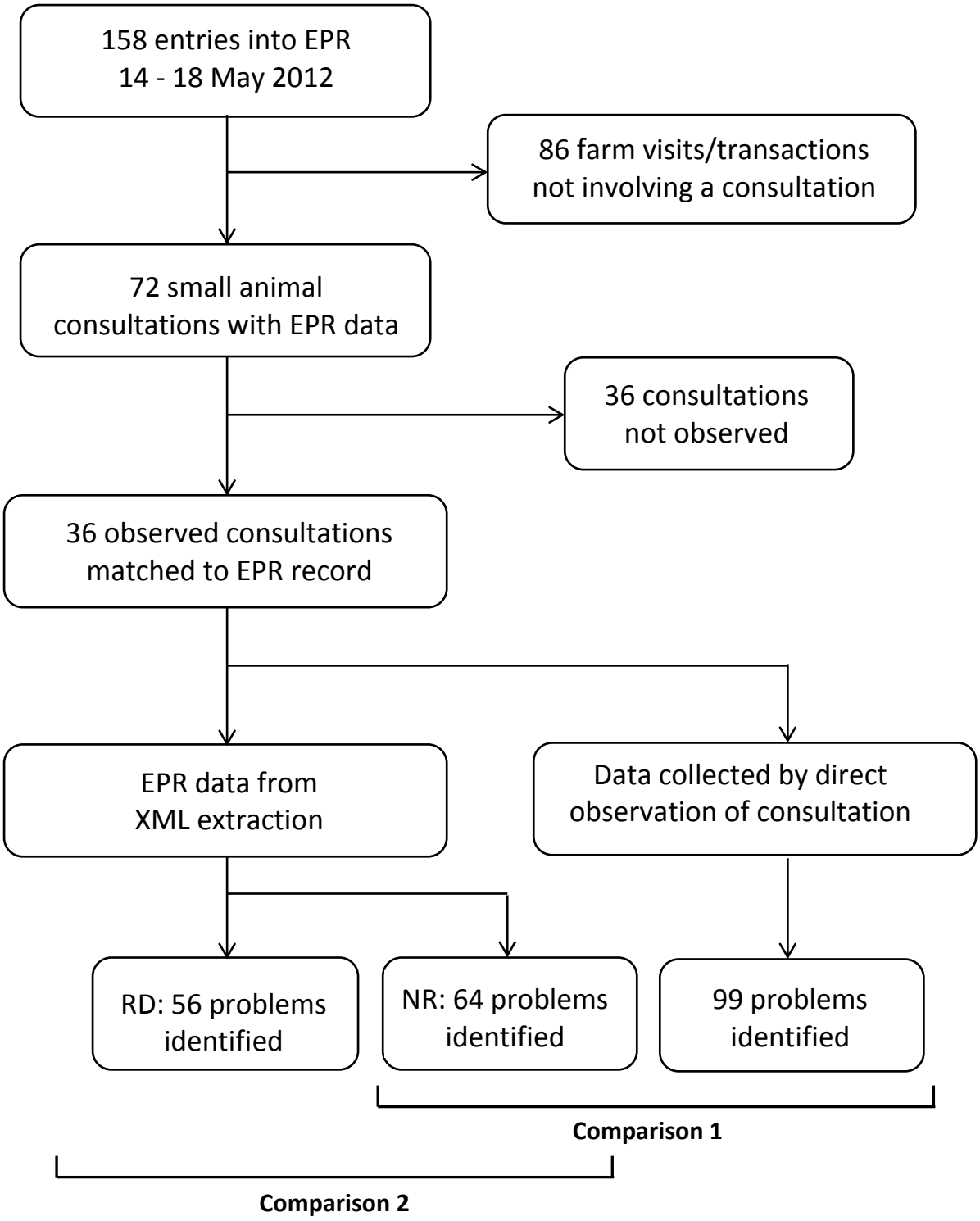

Figure 1. Flow diagram presenting the number of consultation record extracted from the practice computer using the XML schema. The extracted data (n = 158 entries) were then cross-matched with data collected by direct observation (n=36 consultation) to allow examination of the data recorded into the animal record during the observed consultations.

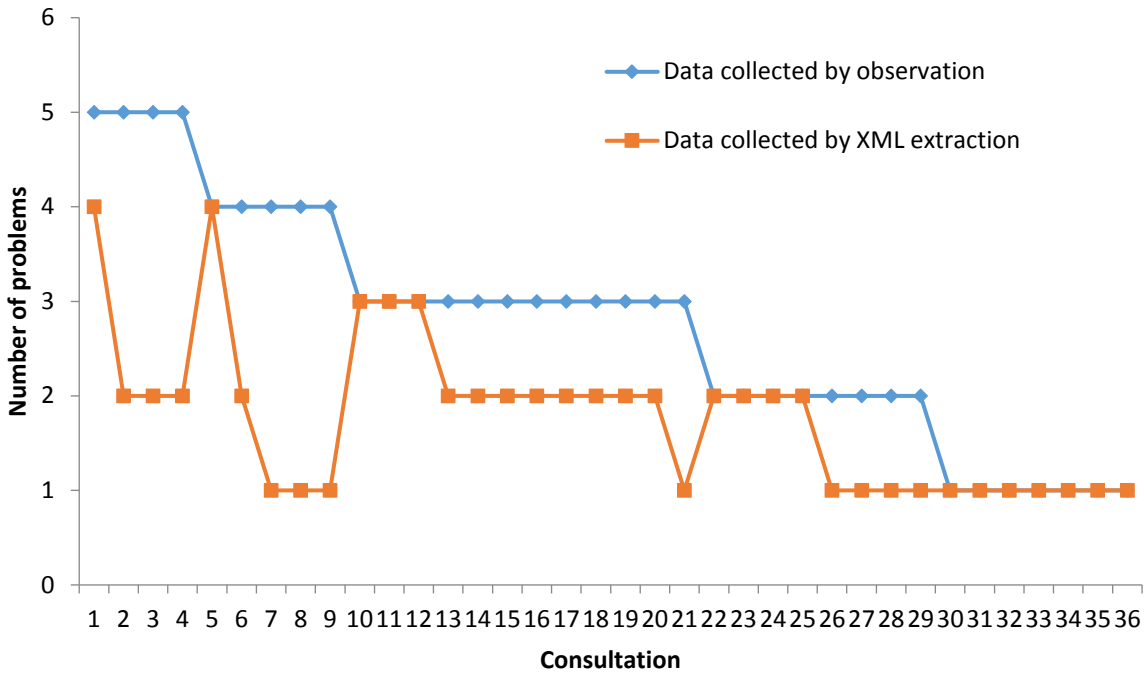

Figure 2. The number of problems observed and recorded during animal consultations using a direct observation method (n = 99 problems) observed by an independent researcher (NR) compared with those problems recorded by the consulting veterinarian and extracted from the practice computer using the XML schema as assessed by the same independent researcher (N=64) for each of the 36 consultations.

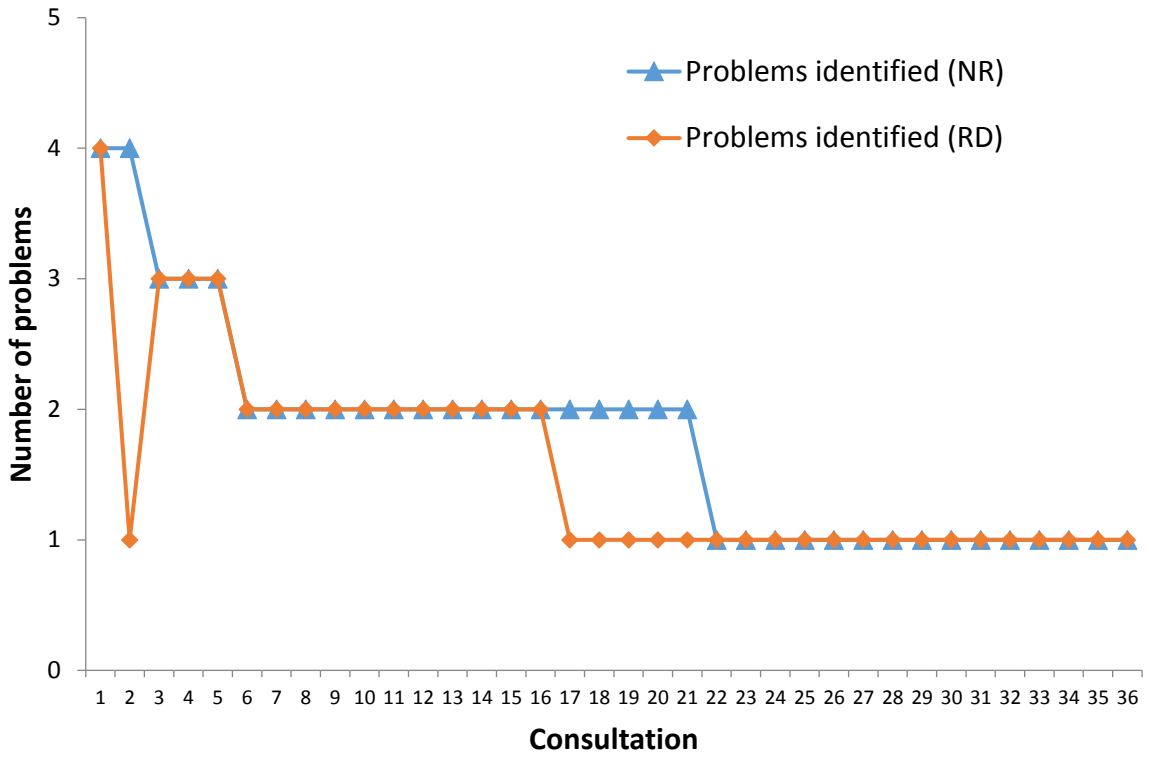

Figure 3. Plot to demonstrate the agreement between two independent reviewers (NR & RD) when comparing the number of problems identified by the researchers (NR = 64, RD = 56) in the XML extracted dataset. The level of agreement was assessed using a linear-weighted Kappa analysis to determine the inter-rater reliability, *K* = 0.727 (95% Confidence Interval 0.502-0.952).