Preventive Veterinary Medicine 94 (2010) 264–271

Contents lists available at ScienceDirect

Preventive Veterinary Medicine

journal homepage: www.elsevier.com/locate/prevetmed

Research paper

Disease surveillance and referral bias in the veterinary medical database

Paul C. Bartlett [a,*], James W. Van Buren [b], Margaret Neterer [c], Chun Zhou [d]

[a] College of Veterinary Medicine, Michigan State University, E. Lansing, MI 48824, United States
[b] Vector Scientific Resources, Inc., 5780 Braeburn Court, Kalamazoo, MI 49009, United States
[c] Michigan State University, E. Lansing, MI 48824, United States
[d] Hill's Pet Nutrition, 400 S W Eighth Avenue, Topeka, KS 66603, United States

ARTICLE INFO

Article history:
Received 26 May 2009
Received in revised form 7 January 2010
Accepted 10 January 2010

Keywords:
Canine
Feline
Epidemiology
Surveillance

ABSTRACT

The Veterinary Medical Database (VMDB) is a summary of veterinary medical records from North American veterinary schools, and is a potential source of disease surveillance information for companion animals. A retrospective record search from four U.S. university veterinary teaching hospitals was used to calculate crude disease rates. Our objectives were to evaluate the utility of the database for disease surveillance purposes by comparing the utility of two methodologies for creating disease categories, and to evaluate the database for evidence of referral bias. Summaries of the medical records from November 2006 to October 2007 for 9577 dogs and 4445 cats were retrieved from VMDB for all canines and felines treated at Kansas State University, Colorado State University, Purdue University and Ohio State University. Disease frequency, computed as apparent period-prevalence and as the percentage of veterinary visits, was compiled for 30 disease categories that were formulated by one of two methods. To assess the possible impact of referral bias, disease rates were compared between animals residing in zip codes within 5 miles of the hospitals (zone 1) and those animals living at more distant locations (zone 2).

When compared to zone 1 animals, disease conditions commonly associated with primary veterinary care were reduced by 29–76% within zone 2, and selected diseases generally associated with more specialized care were increased from 46 to 80% among zone 2 animals. The major differences in disease prevalence seen between zones suggests that substantial referral bias may exist, and that adjustment on the basis of geographical proximity to the university teaching hospitals may be useful in reducing this type of selection bias in the VMDB, thereby improve the accuracy of prevalence estimates and enhancing the utility of this database for purposes of disease surveillance.

1. Introduction

Passively collected disease surveillance data is prone to many types of biases. Nevertheless, field epidemiologists use admittedly imperfect databases to produce estimates of disease occurrence and disease trends that are useful for many purposes. For example, the Morbidity and Mortality Weekly Report (MMWR) is published weekly by the Centers for Disease Control and Prevention, yet multiple types of selection and informational bias heavily influence its disease rates that are based entirely on passively collected observational field data. For the purposes of disease surveillance, a reasonable estimate is often preferable to no estimation at all, even though observation surveillance data is clearly prone to many more biases than would be expected in data collected for a well-designed and controlled laboratory experiment.

Disease surveillance within companion animal populations is useful for many purposes, including veterinary

* Corresponding author. Tel.: +1 517 884 2016; fax: +1 517 432 2310.
E-mail address: Bartlett@cvm.msu.edu (P.C. Bartlett).

0167-5877/$ – see front matter © 2010 Published by Elsevier Ltd.
doi:10.1016/j.prevetmed.2010.01.007

service resource projections, directing teaching curriculums, detecting disease trends and also for animal health product development. Additionally, the 2007 pet food melamine exposures exemplified how monitoring of disease in an animal population might signal an early warning of a human disease health hazard (Brown et al., 2007, pp. 525–531; Puschner et al., 2007, pp. 616–624; Dobson et al., 2008, pp. 251–262). The founders of the Veterinary Medical Data Base (VMDB) realized these and other applications (VMDB, 2008). The VMDB was started in 1964 by the National Cancer Institute to abstract the medical records of the North American veterinary schools (Folk et al., 2002, pp. 405–410; Hahn et al., 2004, pp. 475–479). Currently, 26 veterinary schools have submitted data, but many are many years behind in their data submission. Medical records at each school are transcribed into SNOMED (Systematized Nomenclature of Medicine) codes and then centralized to a national database. Only basic demographics, dates, diagnosis and medical procedures are collected by VMDB. Patient number and owner identifiers link the VMDB database to each hospital's own medical records system, where researchers can retrieve more complete records upon approval by each hospital. However, the VMDB database manager reports that many schools are several years behind in SNOMED coding their medical records, and some institutions contribute data from only some of their clinical services. Nevertheless, the VMDB is the oldest and probably the most comprehensive pet animal health database in the U.S., and its utility for disease surveillance purposes and observational studies warrants evaluation.

Referral bias is a type of selection bias that affects estimates of disease frequency if the monitored clinics or hospitals have a predominance of patients that were referred for more specialized care, and may also exist if they have a relative deficit of cases receiving primary care (Salive, 1994, pp. 808–809; Sepkowitz, 1990, pp. 1629–1631; Froom and Froom, 1992). The extent of referral bias is expectedly different for every disease diagnosis, so no one measure of referral bias would be expected for all diseases or disease categories. Referral bias is, therefore, problematic if one's objective is to estimate disease frequency among the entire segment of the pet population that utilizes both primary and referral veterinary services.

Disease rates are the most basic of all the epidemiologic measures. Most clinicians realize that a clinical or laboratory value, e.g. BUN or body temperature, is useless in the absence of information regarding the expected or normal value in the healthy population. Similarly, disease rates for one particular veterinary clinic have little meaning unless compared to a large reference population. With up to 26 possible participating clinics, the VMDB database has the potential to serve as a reference population for national animal disease rates, if and when referral bias and other biases can be ameliorated. Private practitioners may wish to compare their practice caseloads to this reference population to identify areas for practice expansion or contraction of services, and to identify and monitor unexpected clusters of disease in their particular geographical area.

Our objectives were to evaluate the utility of the database for disease surveillance purposes by comparing the utility of two methodologies for creating disease categories, and to evaluate the database for evidence of referral bias.

2. Methods and materials

2.1. The VMDB database

When schools submit their data to VMDB, they indicated the start date and end date of each data file contribution. A monthly count of records was used to verify that the monthly case count was consistent with previous months. The VMDB database was searched to identify veterinary colleges that were the most current in their data submissions from all of their companion animal medical services, including the 'General Medicine' services that are excluded by some institutions. In September 2008, we searched for the most recent contiguous 12-month period for which we had a complete database for at least four contributing schools. We searched back to November 2006 through October 2007 to meet these conditions, whereupon we selected for study all feline and canine medical records from Kansas State University (KSU), Colorado State University (CSU), Purdue University (PUR) and Ohio State University (OSU). All available data fields were obtained as spreadsheet files, and further data processing was done in C++ and SAS (C++ Resources Network; SAS 9.1.3, 2008).

SNOMED is a hierarchical coding system that includes general disease categories, more specific subcategories and very specific diagnoses. We received assistance from VMDB managers in producing lists of codes regarding 12 general disease categories (Tables 1–4). For these 12 general disease categories, a C++ program was used to identify the absence or presence of at least one of the codes in each of these 12 disease categories. Each of these 12 categories can be identified within column 1 of Table 1 because they all contained over 1000 codes. The additional 18 disease classifications (Tables 1–4) were identified by sorting the spreadsheet files by SNOMED code and searching the database for key words depicting each of these general disease conditions. For example, four codes were found that referred to either obesity or overweight and 28 codes related to kidney diseases (Table 1; column 1). SNOMED codes used for each disease category can be obtained from the authors.

2.2. Analysis by veterinary visit

A SAS program was used to calculate the percentage of veterinary visits (PVV) at which each disease category was diagnosed. For example, a patient would contribute 3 visits to the ear disease numerator and 4 visits to the ear disease denominator if that animal had an ear disease mentioned on 3 of the 4 days that that animal was seen at the clinic during the year. Every day of in-patient care was tabulated as a separate visit.

2.3. Apparent period-prevalence (APP)

For some diseases, veterinarians may be more interested in estimating the percentage of animals for which a

Table 1
Feline cases. Veterinary visits from KSU, CSU, Purdue and OSU—November 2006 through October 2007.

No. of codes	Disease	Zip code within 5 miles				Zip code over 5 miles away			
		Visits	%	Low CI[a]	Up CI	Visits	%	Low CI	Up CI
		4731 (total)				3125 (total)			
3121	Urinary	582	12	11	13	408	13	12	14
12,990	Neurology	234	5.0	4.3	5.6	142	4.5	3.8	5.3
5	Seizures	20	.42	.24	.61	13	.42	.19	.64
39	Liver	114	2.4	2.0	2.9	99	3.2	2.6	3.8*
28	Kidney	361	7.6	6.9	8.4	272	8.7	7.7	9.7
50	Dental	830	18	16	19	192	6.1	5.3	7.0*
1165	Behavioral	67	1.4	1.1	1.8	30	.96	.62	1.3
7234	Neoplasm	292	6.2	5.5	6.9	550	18	16	19*
6621	Eye	539	11	10	12	269	8.6	7.6	9.6*
3	Glaucoma	12	.25	.11	.40	10	.32	.12	.52
2146	Ear	327	6.9	6.2	7.6	160	5.1	4.4	5.9*
10	Otitis	247	5.2	4.6	5.9	108	3.5	2.8	4.1*
4461	Gastrointestinal	316	6.7	6.0	7.4	296	9.5	8.5	11*
2168	Parasitic	74	1.6	1.2	1.9	22	.70	.41	1.0*
5786	Reproductive	18	.38	.21	.56	22	.70	.41	1.0
4	Neutering	231	4.9	4.3	5.5	271	8.7	7.7	9.7*
4	Obesity	270	5.7	5.1	6.4	51	1.6	1.2	2.1*
2	Diabetes	101	2.1	1.7	2.6	68	2.2	1.7	2.7
1	Hypercortisolism	8	.17	.05	.29	0	–	–	–
1	Hypothyroidism	4	.08	.00	.17	6	.19	.04	.35
1	Hyperthyroidism	156	3.3	2.8	3.8	143	4.6	3.8	5.3*
20,866	Musculoskeletal	285	6.0	5.4	6.7	246	7.9	6.9	8.8
2	Hip dysplasia	0	–	–	–	1	.03	.00	.09*
10,033	Dermatologic	754	16	15	17	359	11	10	13*
1	Pyoderma	16	.34	.17	.50	14	.45	.21	.68*
10,113	Cardiovascular	440	9.3	8.5	10	497	16	15	17
2	Heart dx	73	1.5	1.2	1.9	155	5.0	4.2	5.7
3	Recheck	160	3.4	2.9	3.9	127	4.1	3.4	4.8*
8	Vaccination	1226	26	25	27	244	7.8	6.9	8.8*
1	Euthanasia	132	2.8	2.3	3.3	110	3.5	2.9	4.2*

[a] Low CI = lower 95% confidence limit; up CI = upper 95% confidence limit.
* $P < .05$ by Z test for the comparison between two proportions.

particular disease condition was diagnosed at any time during the year. For example, if someone was interested in identifying what percent of the caseload had diabetes, they would want to know the percentage of patients seen during the year that had a diagnosis of diabetes during at least one of those visits. We include the term "apparent" in recognition of the fact that not all dogs or cats seen at least once during the monitored year would necessarily have been presented to this same hospital if and when they developed each of the diseases we studied. Using this analytical approach, each patient seen one or more times during the year contributed to the denominator, and contributed to the numerator only if the particular disease category was diagnosed at least once during the monitored year. This method of analysis is our best estimate of period-prevalence in that it counts any new or continuing cases that existed at any time during the year.

2.4. Stratification by zip code

Zip code provided our only measure of geographical location of the patient. For each school, we identified those zip codes for which at least some of the zip code area was located within 5 miles of the university veterinary clinic (zone 1). Patients were designated as zone 2 if their residence was located in a zip code that was entirely located more than 5 miles from the university veterinary hospital. Our rationale in stratifying on zone was that selected routine primary care diseases such as obesity, dermatologic, vaccination, dental and otitis might be under-represented in the predominantly referral cases from zone 2, and that selected diseases usually requiring specialized veterinary attention (urinary, kidney, liver, cardiovascular and neoplasm) would be over-represented in the largely referral cases living in zone 2. Differences in disease rates between the two zones would be seen as possible evidence of referral bias.

3. Results

Canine patients were 2785 (29%) from CSU, 2158 (23%) from KSU, 2248 (23%) from OSU and 2386 (25%) from PUR. Feline patients were 1590 (36%) from CSU, 1286 (29%) from KSU, 626 (14%) from OSU and 943 (21%) from PUR. Canine patients averaged 6.3 years of age (std = 4.1) and feline patients averaged 5.9 years of age (std = 5.2). Canines from zone 1 averaged 5.8 years of age (std = 4.2) compared to 6.6 yrs of age (std = 4.1) for zone 2 ($P < .0001$ by t-test). Felines from zone 1 averaged 5.4 years of age (std = 4.5) compared to 6.6 yrs of age (std = 5.4) for zone 2 ($P < .0001$ by t-test).

The results of the PVV analysis (Tables 1 and 2) and APP analysis (Tables 3 and 4) are presented for both felines and canines. Major differences between zones 1 and 2 were observed in most disease prevalence rates, as indicated by

Table 2
Canine cases. Veterinary visits from KSU, CSU, Purdue and OSU—November 2006 through October 2007.

Disease	Zip code within 5 miles				Zip code over 5 miles away			
	Visits	%	Low CI[c]	Up CI	Visits	%	Low CI	Up CI
	8788 (total)				13,190 (total)			
Urinary[b]	507	5.8	5.3	6.3	1117	8.5	8.0	8.9*
Neurology	692	7.9	7.3	8.4	1019	7.7	7.3	8.2*
Seizures	139	1.6	1.3	1.8	171	1.3	1.1	1.5
Liver[b]	172	2.0	1.7	2.3	468	3.6	3.2	3.9*
Kidney[b]	136	1.6	1.3	1.8	347	2.6	2.4	2.9*
Dental[a]	1270	14	14	15	484	3.7	3.4	4.0*
Behavioral	68	.77	.59	.96	83	.63	.49	.76
Neoplasm[b]	1363	16	15	16	3674	28	27	29*
Eye	1064	12	11	13	1686	13	12	13
Glaucoma	53	.60	.44	.76	167	1.3	1.1	1.5*
Ear	868	9.9	9.3	11	682	5.2	4.8	5.6*
Otitis[a]	665	7.6	7.0	8.1	477	3.6	3.3	3.9*
Gastrointestinal	679	7.7	7.2	8.3	769	5.8	5.4	6.2*
Parasitic	97	1.1	.89	1.3	74	.56	.43	.69*
Reproductive	90	1.0	.81	1.2	143	1.1	.91	1.3
Neutering	219	2.5	2.2	2.8	220	1.7	1.5	1.9*
Obesity[a]	259	3.0	2.6	3.3	124	.94	.78	1.1*
Diabetes	63	.72	.54	.89	152	1.2	.97	1.3*
Hypercortisolism	61	.69	.52	.87	139	1.1	.88	1.2*
Hypothyroidism	110	1.3	1.0	1.5	123	.93	.77	1.1*
Hyperthyroidism	3	.03	.00	.07	2	.02	.00	.04
Musculoskeletal	1374	16	15	16	2539	19	19	20*
Hip dysplasia	49	.56	.40	.71	131	.99	.82	1.2*
Dermatologic[a]	2094	24	23	25	2193	17	16	17*
Pyoderma	245	2.8	2.4	3.1	286	2.2	1.9	2.4*
Cardiovascular[b]	644	7.3	6.8	7.9	1413	11	10	11*
Heart dx	121	1.4	1.1	1.6	486	3.7	3.4	4.0*
Recheck	619	7.0	6.5	7.6	1216	9.2	8.7	9.7*
Vaccination[a]	1402	16	15	17	500	3.8	3.5	4.1*
Euthanasia	164	1.9	1.6	2.2	273	2.1	1.8	2.3

[a] Selected primary care conditions.
[b] Selected specialized care conditions.
[c] Low Cl = lower 95% confidence limit; up Cl = upper 95% confidence limit.
* $P < .05$ by Z test for the comparison between two proportions.

the confidence intervals for the difference between two proportions (Tables 1–4). Compared to the zone 1 rates, selected canine primary care diseases in zone 2 were decreased by 29–76%, and our five selected diseases thought to require primarily specialized care were increased from 46 to 80% in zone 2 (Figs. 1–4). Other disease categories showed varying differences in disease rates between the two zones.

4. Discussion

If one is interested in using clinical records to measure disease occurrence in the entire pet population, a large selection bias will always exists in that some companion animals almost never receive any veterinary care if and when they become ill. Disease frequency is essentially immeasurable among the segment of the animal population that does not receive veterinary care, except perhaps by costly community-based surveys. Another segment of the pet population may receive primary veterinary care, but their owners are not able to afford the more specialized and expensive treatments offered predominantly at referral clinics. Whether or not a measure of disease occurrence is biased depends on the target population of interest. Disease surveillance at primary care clinics may be viewed as being biased by the relative absence of many serious or rare diseases that are more commonly diagnosed and treated at referral clinics. Disease surveillance at the referral clinics may be viewed as being biased by the over-representation of referral cases requiring specialized veterinary care.

While the case loads at most veterinary schools contain a predominance of referral cases, these university clinics also serve as the primary care veterinary clinics for local residents. It is our experience that fees for routine medical issues are structured to be comparable with local private veterinary clinics so that local pet owners are encouraged to submit routine medical issues to provide teaching material for the veterinary students. Therefore, pet owners residing in the area immediately around veterinary teaching hospitals are able to utilize the university clinics for their primary veterinary medical care and also use this same university hospital for specialized veterinary care if and when it is medically indicated. As such, the medical records of animals residing near large university veterinary clinics may provide a somewhat more realistic balance of both primary care and referral cases compared to the entire university caseload. Residual referral bias may still exist within the zone 1 if clients living in zone 1 utilize a private veterinary clinic for at least some of their

Table 3
Feline cases. Apparent period-prevalence from KSU, CSU, Purdue and OSU—November 2006 through October 2007.

Disease	Zip code within 5 miles				Zip code over 5 miles away			
	Cats 2557 (total)	%	Low CI[a]	Up CI	Cats 1888 (total)	%	Low CI	Up CI
Urinary	310	12	11	13	264	14	12	16
Neurology	186	7.3	6.3	8.3	98	5.2	4.2	6.2*
Seizures	16	.63	.32	.93	12	.64	.28	.99
Liver	55	2.2	1.6	2.7	75	4.0	3.1	4.9*
Kidney	150	5.9	5.0	6.8	160	8.5	7.2	9.7*
Dental	591	23	21	25	146	7.7	6.5	8.9*
Behavioral	62	2.4	1.8	3.0	28	1.5	.94	2.0*
Neoplasm	121	4.7	3.9	5.6	251	13	12	15*
Eye	348	14	12	15	179	9.5	8.2	11*
Glaucoma	7	.27	.07	.48	4	.21	.00	.42
Ear	217	8.5	7.4	9.6	105	5.6	4.5	6.6*
Otitis	159	6.2	5.3	7.2	70	3.7	2.9	4.6*
Gastrointestinal	207	8.1	7.0	9.2	191	10	8.8	11*
Parasitic	71	2.8	2.1	3.4	21	1.1	.64	1.6*
Reproductive	15	.59	.29	.88	22	1.2	.68	1.7*
Neutering	230	9.0	7.9	10	271	14	13	16*
Obesity	230	9.0	7.9	10	41	2.2	1.5	2.8*
Diabetes	29	1.1	.72	1.5	28	1.5	.94	2.0
Hypercortisolism	1	.04	.00	.12	0	–	–	–
Hypothyroidism	2	.08	.00	.19	4	.21	.00	.42
Hyperthyroidism	67	2.6	2.0	3.2	100	5.3	4.3	6.3*
Musculoskeletal	213	8.3	7.3	9.4	169	9.0	7.7	10
Hip dysplasia	0	–	–	–	1	.05	.00	.16
Dermatologic	512	20	18	22	229	12	11	14*
Pyoderma	9	.35	.12	.58	11	.58	.24	.93
Cardiovascular	299	12	10	13	355	19	17	21*
Heart dx	68	2.7	2.0	3.3	141	7.5	6.3	8.7*
Recheck	95	3.7	3.0	4.5	77	4.1	3.2	5.0
Vaccination	1032	40	38	42	215	11	10	13*
Euthanasia	132	5.2	4.3	6.0	110	5.8	4.8	6.9

[a] Low CI = lower 95% confidence limit; up CI = upper 95% confidence limit.
* $P < .05$ by Z test for the comparison between two proportions.

veterinary needs. Zip code was the only geographic information available to us on the VMDB, so a quantitative adjustment for geographic distance was impossible.

Other supporting evidence for referral bias in the VMDB was the preponderance of older animals in zone 2. This predominance of older animals in referral case populations is consistent with age distributions found in other animal health databases (Antech FHT Study, 2006; Hills Consumer Perceived Condition Study, 2008).

One limitation of using general disease categories for displaying disease surveillance information is that any increase in only one code may be difficult to detect. For example, a several fold increase in acute renal failure, such as that due to the recent pet food Melamine outbreak, may go unrealized due to the many other codes contained in the same renal category. In this regard, automated timely data-mining programs might be utilized to detect those specific SNOMED codes that are substantially increased above a rolling baseline average. If dramatic increases in any one code are detected, the VMDB database offers an excellent opportunity to identify cases for follow-up case–control studies (Ward, 2002, pp. 203–213; Gelatt and MacKay, 2004a, pp. 97–111; Gelatt and MacKay, 2004b, pp. 245–259; Bryan et al., 2007, pp. 1174–1181). The 12 disease categories based on the SNOMED hierarchical structure each contained over 1000 SNOMED codes, so it is difficult to believe that an increase in any one of these codes would not be completely diluted out by being combined with so many other codes.

One of the advantages in using the VMDB database for disease surveillance is the relative reliability of the diagnoses. For example, few veterinary college clinicians will diagnose chronic renal failure without laboratory support as well as appropriate clinical signs. In contrast, disease surveillance systems from primary care clinics may record initial presenting complaints rather than the final veterinary diagnosis. One of us (JB) has done extensive work with collecting disease surveillance data from veterinary clinic management software, and learned that most computer software was designed principally for business management, and diagnoses may be recorded as very broad disease categories without any potential for more specific or refined data retrieval (Lund et al., 1999, pp. 1336–1341).

Both calculation methods (PVV and APP) include new and continuing cases. A disadvantage of the APP method is that it gives equal weight in the database to an animal that was seen once during the year and another animal that was seen multiple times during the year. The APP also presumes that each animal seen at least once during the year (and therefore contributing to the denominator) would necessarily be presented to this teaching hospital for each and every disease if and when it occurred. This certainly is not always true, as many clients taking their animals to the university hospital for some procedures may utilize the services of a different clinic for their other veterinary needs. The PVV method of analysis may be the

Table 4
Canine cases. Apparent period-prevalence from KSU, CSU, Purdue and OSU—November 2006 through October 2007.

Disease	Zip code within 5 miles				Zip code over 5 miles away			
	Dogs 3502 (total)	%	Low CI[a]	Up CI	Dogs 6075 (total)	%	Low CI	Up CI
Urinary	278	7.9	7.0	8.8	630	10	9.6	11*
Neurology	436	12	11	14	695	11	11	12
Seizures	62	1.8	1.3	2.2	109	1.8	1.5	2.1
Liver	122	3.5	2.9	4.1	333	5.5	4.9	6.1*
Kidney	64	1.8	1.4	2.3	198	3.3	2.8	3.7*
Dental	824	24	22	25	349	5.7	5.2	6.3*
Behavioral	64	1.8	1.4	2.3	74	1.2	.94	1.5*
Neoplasm	526	15	14	16	1275	21	20	22*
Eye	669	19	18	20	971	16	15	17*
Glaucoma	20	.57	.32	.82	85	1.4	1.1	1.7*
Ear	521	15	14	16	431	7.1	6.5	7.7*
Otitis	391	11	10	12	301	5.0	4.4	5.5*
Gastrointestinal	479	14	13	15	518	8.5	7.8	9.2*
Parasitic	81	2.3	1.8	2.8	58	.95	.71	1.2*
Reproductive	66	1.9	1.4	2.3	109	1.8	1.5	2.1
Neutering	212	6.1	5.3	6.8	216	3.6	3.1	4.0*
Obesity	200	5.7	4.9	6.5	97	1.6	1.3	1.9*
Diabetes	19	.54	.30	.79	53	.87	.64	1.1
Hypercortisolism	28	.80	.50	1.1	65	1.1	.81	1.3
Hypothyroidism	66	1.9	1.4	2.3	85	1.4	1.1	1.7
Hyperthyroidism	2	.06	.00	.14	2	.03	.00	.08
Musculoskeletal	768	22	21	23	1485	24	23	26*
Hip dysplasia	26	.74	.46	1.0	102	1.7	1.4	2.0*
Dermatologic	1116	32	30	33	1183	19	18	20*
Pyoderma	160	4.6	3.9	5.3	191	3.1	2.7	3.6*
Cardiovascular	350	10	9.0	11	882	15	14	15*
Heart dx	88	2.5	2.0	3.0	396	6.5	5.9	7.1*
Euthanasia	164	4.7	4.0	5.4	273	4.5	4.0	5.0
None of above	608	17	16	19	732	12	11	13*
Recheck	348	9.9	9.0	11	717	12	11	13*
Vaccination	1138	33	31	34	419	6.9	6.3	7.5*

[a] Low Cl = lower 95% confidence limit; up Cl = upper 95% confidence limit.
* $P < .05$ by Z test for the comparison between two proportions.

most useful for veterinary practitioners seeking a basis of comparison for their practice's caseload. However, this method of calculation may be less useful for persons seeking information on the occurrence of a chronic debilitating disease condition within an animal population.

The magnitude of the disease rate differences between zones 1 and 2 is particularly striking for some disease categories. Differences exceed 75% for some diseases, and suggest major differences in utilization of veterinary services between zones 1 and 2. Because of the magnitude of these differences for some diseases, issues of referral bias must be considered in any estimates of disease frequency utilizing the VMDB database. Even within zone 1, residual referral bias may occur in that pet owners may utilize local private clinics for routine matters, but take

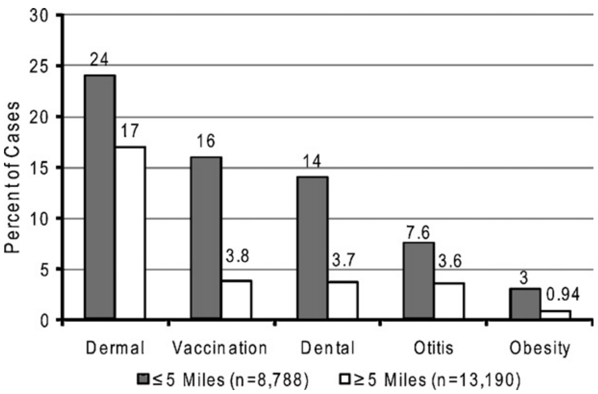

Fig. 1. Primary care canine diseases for KSU, CSU, Purdue and OSU, November 2006–October 2007.

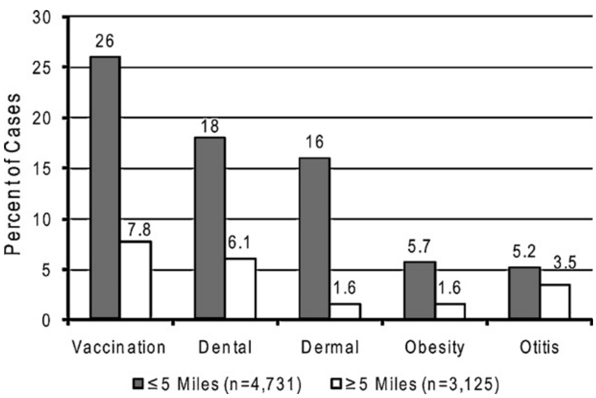

Fig. 2. Primary care feline diseases for KSU, CSU, Purdue and OSU, November 2006–October 2007.

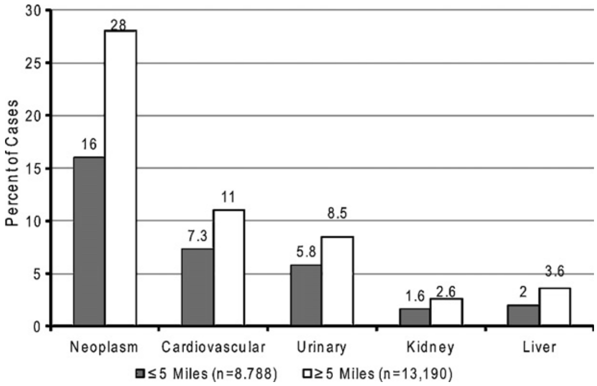

Fig. 3. Referral care canine diseases for KSU, CSU, Purdue and OSU, November 2006–October 2007.

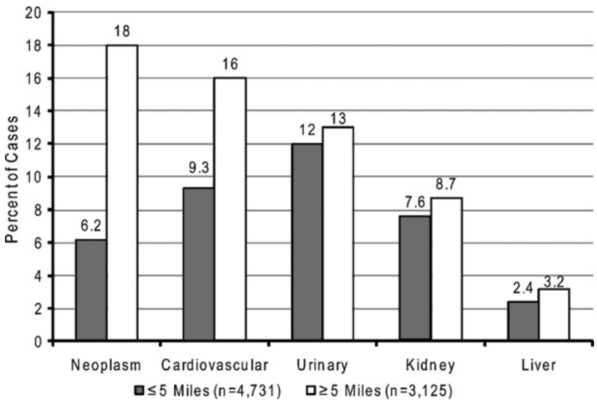

Fig. 4. Referral care feline diseases for KSU, CSU, Purdue and OSU, November 2006–October 2007.

their pet to the university hospital when more specialized care is needed. Therefore, differences between zones 1 and 2 may be a minimum estimate of the true referral bias that might exist actually exist.

Controlling or adjusting for referral bias is beyond the scope of the current study. Restricting the database to nearly localities greatly reduces the size of the database, but may reduce the extent of referral bias. However, our zone 1 confidence intervals are sufficiently small for most disease surveillance purposes, and there is potential to reduce these confidence intervals by increasing the size of the database by the addition of more universities or more years of collected data. Unless one is only interested in disease trends, a less biased estimate is usually preferred over a biased estimate for most disease surveillance purposes, even at the expense of some precision. If geographic distance could be quantified for each case, multivariable adjustment for geographical distance might be employed to help adjust for referral bias.

Whether a bias is reduced by matching, statistical adjustment or stratification, the extent of any bias can only be estimated by comparing the biased estimate to a less biased estimate. In the current study, stratification by zone revealed substantial differences between zones for several diseases, suggesting that referral bias may be substantial. If it can be assumed that the caseload within zone 1 is reasonably representative of both primary care and secondary care, then our disease estimates for zone 1 are likely to be less affected by referral bias and, therefore, more accurate than are disease rates shown for zone 2.

Few other authors have reported disease surveillance information for the U.S. companion animal population. Using data from 1995, the Lund et al. (1999) study utilized the records of 52 volunteering veterinary practices that all used the same practice management software. Due to irreconcilable differences in disease category definitions, our estimates of disease prevalence cannot be compared with those made by Lund.

The SNOMED hierarchical coding scheme was developed primarily for human medicine. Use of this hierarchical organizational structure to identify 12 broad categories proved unnecessarily cumbersome because the medical records coders at the schools tended to repeatedly use a much smaller subset of the available SNOMED codes. Because the VMDB database contained a short description of the meaning of each SNOMED code, it was easy and transparent to manually create disease classification categories by searching for key words indicative of each major disease condition. The resultant disease categories could then be easily subdivided into specific SNOMED codes for a more detailed analysis. Our subjective assessment was that it was easier, more specific, more meaningful and more useful to use our 18 manually created disease categories rather than the 12 categories created from the SNOMED hierarchical system.

5. Conclusion

In summary, the VMDB has some potential for estimating rates of canine and feline disease occurrence. For some diagnostic categories, we found evidence of major differences in disease rates when local patients were compared to those residing in more distant locations, suggesting that major referral bias may be inherent in the database for some diagnoses. The VMDB may ultimately prove valuable as a source of disease surveillance information if analytical methods can be developed to reduce or adjust for the inherent referral bias. Two different methods for computing disease rates each had their advantages and disadvantages, but we believe the PPV is more applicable to the needs of most clinicians. Especially if standardized veterinary disease coding groups could be developed, the VMDB system could become more valuable for disease surveillance purposes if the veterinary schools more fully participate by contributing data from all their clinical services and by submitting their data in a timely fashion.

Acknowledgement

Funding for this project was provided by Hills Pet Nutrition.

References

Brown, C.A., Jeong, K.S., Poppenga, R.H., Puschner, B., Miller, D.M., Ellis, A.E., Kang, K.I., Sum, S., Cistola, A.M., Brown, S.A., 2007. Outbreaks of renal failure associated with melamine and cyanuric acid in dogs and cats in 2004 and 2007. J. Vet. Diagn. Invest. 19 (5), 525–531.

Bryan, J.N., Keeler, M.R., Henry, C.J., Bryan, M.E., Hahn, A.W., Caldwell, C.W., 2007. A population study of neutering status as a risk factor for canine prostrate cancer. Prostrate 67 (11), 1174–1181.

C++ Resources Network. www.Cplusplus.com.

Dobson, R.L., Motlagh, S., Quijano, M., Cambron, R.T., Baker, T.R., Pullen, A.M., Regg, B.T., Bigalow-Kern, A.S., Vennard, T., Fix, A., Reimschuessel, R., Overmann, G., Shan, Y., Daston, G.P., 2008. Identification and characterization of toxicity of contaminants in pet food leading to an outbreak of renal toxicity in cats and dogs. Toxicol. Sci. 106 (1), 251–262.

Folk, L.C., Hahn, A.W., Patrick, T.B., Allen, G.K., Smith, A.B., Wilcke, J.R., 2002. Salvaging legacy data: mapping an obsolete medical nomenclature to a modern one. Biomed. Sci. Instrum. 38, 405–410.

Froom, P., Froom, J., 1992. Selection bias in using data from one population to another: common pitfalls in the interpretation of medical literature. Theor. Med. 13 (3), 255–259.

Gelatt, K.N., MacKay, E.O., 2004a. Prevalence of the breed-related glaucomas in pure-bred dogs in North America. Vet. Ophthalmol. 7 (2), 97–111.

Gelatt, K.N., MacKay, E.O., 2004b. Secondary glaucomas in the dog in North America. Vet. Ophthalmol. 7 (4), 245–259.

Hahn, A.W., Martin, M.K., Siegel, A.M., Ellis, W.K., 2004. Sending data to a central repository. Biomed. Sci. Instrum. 40, 475–479.

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. JAVMA 214 (9), 1336–1341.

Morbidity and Mortality Weekly Report. www.cdc.gov/mmwr.

Puschner, B., Poppenga, R.H., Lowenstine, L.J., Filigenzi, M.S., Pesavento, P.A., 2007. Assessment of melamine and cyanuric acid toxicity in cats. J. Vet. Diagn. Invest. 19 (6), 616–624.

Salive, M.E., 1994. Referral bias in tertiary care: the utility of clinical epidemiology. Mayo Clin. Proc. 69 (8), 808–809.

SAS 9.1.3, 2008. SAS Institute, Raleigh, North Carolina.

Sepkowitz, S., 1990. Effects of selection bias on outcomes at referral centers. Am. J. Obstet. Gynecol. 162 (6), 1629–1631.

VMDB, 2008. http://www.vmdb.org/ (accessed February 2, 2009).

Ward, M.P., 2002. Seasonality of canine leptospirosis in the United States and Canada and its association with rainfall. Prev. Vet. Med. 56 (3), 203–213.