
Ahmed S. Mohamed\textsuperscript{a}; Michael Levine\textsuperscript{b}; Joseph W. Camp Jr.\textsuperscript{a}; Elisabeth Lund\textsuperscript{c} Jonathan S. Yoder\textsuperscript{d}; Larry T. Glickman\textsuperscript{c}; George E. Moore\textsuperscript{a}\textsuperscript{*}

\textsuperscript{a}Department of Comparative Pathobiology, Purdue University, West Lafayette, IN, USA

\textsuperscript{b}Department of Statistics, Purdue University, West Lafayette, IN, USA

\textsuperscript{c}Banfield Pet Hospital\textsuperscript{TM}, Portland, OR, USA

\textsuperscript{d}Division of Foodborne, Waterborne, and Environmental Diseases, National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention (CDC), Atlanta, GA, USA.

\textsuperscript{e}Department of Emergency Medicine, University of North Carolina, Chapel Hill, NC, USA.

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Corresponding author:

Dr. George E. Moore

725 Harrison Street

West Lafayette, IN 47907-2027

Email: gemoore@purdue.edu

Phone: 765-496-3393

Fax: 765-496-2627
Running title: *Giardia* Infection in Human and Canine

Keywords: dogs, *Giardia* spp., infectious disease, protozoa, temporal, surveillance, zoonosis
Giardia protozoa have been suspected of zoonotic transmission, including transmission from companion animals such as pet dogs to humans. Patterns of infection have been previously described for dogs and humans, but such investigations have used different time periods and locations for these two species. Our objective was to describe and compare the overall trend and seasonality of Giardia species infection among dogs and humans in the United States from 2003 through 2009 using public health surveillance data and medical records of pet dogs visiting a large nationwide private veterinary hospital. Canine data were obtained from all dogs visiting Banfield hospitals in the United States with fecal test results for Giardia species, from January 2003 through December 2009. Incidence data of human cases from the same time period was obtained from the CDC. Descriptive time plots, seasonal trend decomposition (STL), and seasonal autoregressive moving-average (SARIMA) model were used to assess the temporal characteristics of Giardia infection in the two species. Canine incidence showed a gradual decline from 2003 to 2009 with a non-significant irregular seasonal component. By contrast, human incidence showed a stable trend with a significant regular seasonal cycle, peaking in August and September. Different temporal patterns in human and canine Giardia cases observed in this study suggest that the epidemiological disease processes underlying both series might be different, and Giardia transmission from dogs to humans and from humans to dogs might be uncommon.
1. Introduction

*Giardia* protozoal parasites infect many species of domestic and wild animals as well as humans. Zoonotic transmission of some *Giardia* species/genotypes has been demonstrated experimentally, but its occurrence and clinical significance under natural conditions is unclear (Plutzer et al., 2010). Assemblages A and B which were considered to be human-specific have been isolated from a wide range of domestic, wild, and marine animals (Thompson et al., 2000), and these zoonotic assemblages have been shown to occur more commonly in dogs from the western United States compared to dog-specific assemblages (C and D) (Covacin et al., 2011). However, the relative importance of zoonotic transmission of *Giardia* spp. remains to be determined (Hunter and Thompson, 2005).

Human giardiasis in the United States is a nationally reportable disease in most states (Yoder et al., 2010). Approximately 20,000 human giardiasis cases were reported annually to the Centers for Disease Control and Prevention (CDC) from 2002 to 2009 (Yoder and Beach, 2007; Yoder et al., 2010), but CDC estimates the actual number of cases to be closer to 1.2 million cases per year due to underreporting and underdiagnosis (Scallan et al., 2011). Documented human giardiasis has been associated with a history of travel, outdoor recreational activities, and drinking contaminated water (Eisenstein et al., 2008).

Cases of human giardiasis in the United States generally increase in late summer and early fall (Katz et al., 2006; Nakada et al., 2012; Yoder et al., 2010). The peak incidence of human giardiasis occurs during the spring in Europe and summer in Canada and the UK (Lal et al., 2012). The seasonality of canine giardiasis has been the subject of conflicting findings. For
example, no seasonal pattern of canine giardiasis in the US was found in one study (Nolan and Smith, 1995), whereas a more recent study reported a highest prevalence in the month of November (Mohamed et al., 2013). The peak incidence of canine giardiasis has been reported to occur in the winter in Italy (Bianciardi et al., 2004), the summer in Spain (Díaz et al., 1996), and in the fall in Argentina (Fontanarrosa et al., 2006).

Time-series analysis is a method for describing the occurrence of common events over time while accounting for the serial correlation (autocorrelation) between observations. Few studies have used a time-series approach to describe the temporal pattern of Giardia (Naumova et al., 2000; Nolan and Smith, 1995). No studies however have compared the temporal patterns of Giardia infections across animal species. Similarities in temporal patterns could potentially indicate common source etiologies or cross-species transmission.

The objective of this study therefore was to describe temporal pattern of giardiasis among dogs and humans in the United States using medical records of dogs visiting private veterinary hospitals and reports of human giardiasis by state health departments to the CDC for the period from January 2003 through December 2009.

2. Materials and methods

2.1. Data:

Canine: Fecal test information was obtained from Banfield, The Pet Hospital, Portland, OR. Fecal testing was performed as part of routine diagnostic or preventive veterinary care of symptomatic and asymptomatic pet dogs during visits to Banfield veterinary hospitals. Fecal
flotation without centrifugation using 1.18 SG ZnSO₄ was performed to detect Giardia cysts in
the stool and the results reported as positive or negative; no attempt was made to identify specific
Giardia assemblages. All fecal tests were conducted by trained hospital staff following using a
standard protocol. The medical records from all Banfield hospitals nationwide are downloaded
weekly and stored in central electronic data warehouse using proprietary software (PetWare,
Banfield, The Pet Hospital, Portland, OR). Each record includes a unique patient and hospital
identifier. Demographic data for each dog including hospital visit date and the results of fecal
flotation tests from January 1, 2003, through December 31, 2009, were downloaded from the
central database. Only results from the first fecal test for each dog were used in the analysis.
Data related to clinical signs if present and specific treatments were not available

The main dataset for canine data was organized into a subset containing all positive fecal test
results indexed by the test date and a second full set containing all fecal tests (positive +
negative) indexed by the test date. A total count of the number of observations in the subset and
the full set was calculated for each month of the seven years, and a monthly incidence ($MP_{d,i}$)
per 100 dogs was calculated as the number of positive fecal tests (in the subset) for each month $i$
($NPT_i$) divided by the total number of tests (in the full set) for the same month ($TNT_i$):

$$MP_{d,i} = \left( \frac{NPT_i}{TNT_i} \right) \ast 100$$

The number of human Giardia cases reported to CDC’s National Notifiable Disease Surveillance
System from each state by month ($TNRC_i$) from January 2003 through December 2009 was
obtained from CDC. An estimate of the total population for each state included in the study for
each of the seven years was obtained from the federal census website (US Census Bureau, 2009).
The total population \((TP_i)\) for each state was used as the denominator to calculate a monthly incidence of Giardia \((MP_{h,i})\) per 100,000 people:

\[
MP_{h,i} = \left( \frac{TNRC_i}{TP_i} \right) \times 100,000
\]

2.2. Analysis:

Monthly incidence rates of canine and human Giardia infection were graphed. The seasonal-trend decomposition procedure based on loess (STL) method (Cleveland et al., 1990; Barnett and Dobson, 2010) was then used to decompose the time series in order to visualize patterns. This procedure is based on decomposing the full time-series into trend, seasonal, and remainder components using a sequence of applications of the local linear regression method (loess) smoother. Additionally monthly data, e.g. all January data, was plotted as a cycle-subseries of the seasonal component.

Model fitting: A mixed Box-Jenkins approach was used to construct appropriate models to describe time-series of Giardia infections in humans and dogs. A script, auto.arima in the package Forecast in R (Hyndman and Khandakar, 2008), was used to produce an initial model which was then refined using a seasonal autoregressive integrated moving average (SARIMA) package (Shumway et al., 2011). SARIMA is an extension of the autoregressive integrated moving average (ARIMA) models and is used to model time series with component(s) that repeat regularly every “S” period of time (seasonal). Accordingly, SARIMA includes seasonal and non-seasonal components; candidate models are first selected on the basis of the exploratory analysis that takes into consideration the time plot structure, properties of the model residuals’ autocorrelation (ACF) and partial autocorrelation (PACF) plots for each series. In the final step,
diagnostics of the residuals and Akaike information criterion (AIC) values were used to select the final model that best fit the data and appeared to satisfy statistical assumptions. In the initial automatic script, integration order, autoregressive (AR), and moving-average (MA) coefficients were selected based on minimizing AIC for the seasonal and non-seasonal components of the model. All statistical analyses were conducted using R (R Development Core Team, 2012) and a p-value <0.05 was considered statistically significant.

3. Results:

A total of 135,802 cases of human giardiasis were reported to the CDC during the period from 2003 through 2009 from all states except IN, KY, NK, MS, and TX where notification was not required during this period. The total annual number of reported human cases remained relatively stable (20,751 in 2004 to 18,478 in 2009) whereas the total number of cases reported from month to month fluctuated with noticeable increases during late summer and early fall. Reported human cases were generally lowest in February (1,216) and highest in August (2,383) during the study period.

Using the first fecal test available in the medical record for each dog, the total number of canine fecal tests included in the study was 2,468,359. These tests were obtained from 777 Banfield veterinary hospitals in 43 states. The number of canine fecal samples tested annually for *Giardia* increased from 288,803 in 2003 to 483,016 in 2009, concurrent with increasing number of hospitals. Meanwhile, the percentage of dogs testing positive for *Giardia* infection annually declined during the study period ranging from 0.61% (1,760/288,803) in 2003 to 0.27% (1,326/483,016) in 2009. There was slight monthly variation in percentage of tests positive from
a low of 0.39% (725/185,371) in November to a high of 0.52% (1,108/214,466) in January. Time
plots of monthly incidence of human giardiasis (per 100,000) and monthly incidence (per 100) of
canine positive fecal tests are shown in Figure 1A and B, respectively.

Examination of human case data with the STL method showed no clear overall trend over time
as the incidence was relatively constant in the calculated trend throughout the seven-year study
period, from a high of 0.75 (cases per 100,000) in 2005 to 0.70 in 2009 (Figure 2A). A regular
seasonal pattern was noticeable with a large magnitude of variation (approximately 0.4
cases/100,000) peaking in July through October (Figure 2B). In contrast, the canine series trend
indicated a general decline over study period from a high of 0.70 (per 100 dogs) in 2003 to 0.30
(per 100 dogs) by the end of 2009 (Figure 3A). The seasonal pattern for this data series however
was irregular with a small magnitude of variation (±0.03 cases/100) (Figure 3B).

Fitting a SARIMA model for the human series indicated a non-seasonal moving-average and an
annual integrated seasonal moving-average term; both terms were statistically significant (Table
1). In comparison, the canine model included an integrated non-seasonal moving average and a
4-month seasonal moving-average term, yet only the non-seasonal term was statistically
significant. Both models were deemed to adequately fit the data given the uncorrelated residuals
and that Ljung-Box Q test of the residuals was not significant.

Note that it is also possible to fit a lagged regression type model where the human incidence
series is regressed against the lags of canine incidence series. Such a model, often called the
lagged regression or a SARIMA(X) model, needs to be approached carefully. As a first
step, the number of lags of canine incidence series to be included has to be estimated. To do this, the so-called prewhitening (see, e.g. Shumway and Stoffer (2011)) has to be applied to both sides of the regression equation. This operation transforms the input (canine giardiasis) series into the white noise and, then, the cross-correlation between the transformed output (human giardiasis) series and the just mentioned white noise. In our specific case, the resulting cross-correlation did not have a single significant lag which indicated that there is very little, if any, dependence between the temporal/seasonal patterns of human and canine giardiasis series.

4. Discussion
To the best of our knowledge, this is the first study to use time-series techniques to analyze temporal patterns of Giardia infection among dogs and people over multiple years in the United States. Human data compared with canine data seemed to follow two different temporal patterns suggesting that the generating processes underlying both series might be different, and that Giardia transmission from dogs to humans and from humans to dogs might be uncommon. The human data series exhibited a strong and regular annual cycle with peaks observed in July through October months, but the canine series did not demonstrate any clearly defined seasonality. Further research will be needed to determine if humans and dogs are simply infected by different assemblages, or if the same assemblage/organism comes from different sources.

The seasonal pattern of human giardiasis observed here is in agreement with prior CDC reports and other studies that reported peaks of human giardiasis in the late summer and early fall.
(Furness et al., 2000; Naumova et al., 2000). Although the main risk factors for giardiasis in humans include contaminated water or food, the increased incidence during late summer months may be attributable to increased human outdoor activities resulting in increased exposures. Interestingly, the human incidence remained relatively constant throughout the seven-year period based on national data. Other trends may have occurred at the state level, but this was not investigated in this study. The ‘stable’ incidence in people may indicate the need for increased efforts to educate the public about potential infection sources and appropriate preventive measures.

The canine data showed a marked downward trend over the study period despite an increased number of dogs being tested at Banfield veterinary hospitals for intestinal parasitism. This trend is unlikely to reflect changes in diagnostic methods as all samples were examined by trained staff following a standardized protocol in all Banfield hospitals. As previous research by our group documented a higher risk of *Giardia* infection in pure breed vs. mixed breed dogs and in younger vs. older dogs (Mohamed et al., 2013), the decreasing prevalence of *Giardia* infection in dogs over time in this study may suggest that fewer puppies are coming from large puppy mills where the prevalence of intestinal parasites is often higher compared with puppies that come from private homes or non-commercial breeders with less crowded or stressful settings (Barr and Bowman, 1994). Alternatively, these sources may be employing more methods in treatment or prevention. The lack of seasonality in the canine series is not totally unexpected, however, and is in agreement with the only available study that analyzed *Giardia* infection among dogs using time-series techniques – albeit at a single hospital location (Nolan and Smith, 1995). Although an
earlier study reported some seasonal patterns (Kirkpatrick, 1988), less rigorous analytic methods in a smaller population were used.

Time-series techniques such as the ones used in this study are useful to analyze and interpret temporal patterns of infection observed using routinely collected surveillance or hospital data. These methods are commonly used in fields such as econometrics, but their application to veterinary medical data has been limited (Benschop et al., 2008; Sanchez-Vazquez et al., 2012). A recent paper (Christiansen et al., 2012) that reviewed methods used to assess seasonality in epidemiological studies of human infectious disease did not include methods such as STL and ARIMA/SARIMA.

STL techniques provide an effective tool to visualize and explore time-series events by dividing them into trend, seasonal, and remainders components that best fit the data (Cleveland et al., 1990). Other methods used to analyze epidemiological data collected over time include generalized linear models (GLM) focusing on evaluating change-point of time parameters rather than decomposing and describing its elements (Christiansen et al., 2012). The SARIMA approach is another equally effective time-series analysis method (Jiang et al., 2010) and was used here to provide some contrast and to verify results obtained from descriptive analyses, i.e. STL. The SARIMA model confirmed that human series follow an annual cycle with a highly significant seasonal component whereas seasonality of the canine series was rather weak. These various techniques are available and can be incorporated in epidemiologic analysis using statistical software such as R (R Development Core Team, 2012).
It is important to point out that both data sets have some shortfalls that could limit the scope of interpreting the observed results. The human data was based on passive surveillance of *Giardia* infection which is believed to be highly underreported (Nakada et al., 2012). The canine data by comparison, despite the exceptionally large sample size, was based on the routinely performed fecal flotation testing which is less sensitive compared to other diagnostic techniques such as centrifugal flotation and ELISA (Zajac et al., 2002; Dryden et al., 2006). Additionally, the fecal *Giardia* test results from dogs did not distinguish whether the dog being tested was asymptomatic and the test was part of a routine wellness exam, or whether it was showing clinical signs associated with an intestinal illness. In contrast, a higher proportion of the human fecal tests were probably performed on individuals who were clinically symptomatic at the time. Due to its retrospective nature, this study was limited in its capability to assess zoonotic risk or source of infection in either species. Ideally, these would be evaluated by performing fecal tests on dogs and humans in the same household at the same time.

5. Conclusion:

Time-series analysis of *Giardia* infection among humans and dogs in the United States for the period from 2003 through 2009 showed that the temporal characteristics of the two data series were different. The human data series exhibited a strong annual seasonal cycle, peaking in August and September, and overall maintained a relatively constant incidence level during the study period. The canine series over the same seven-year period had weak and irregular seasonal fluctuation with an overall declining incidence trend. These findings suggest that underlying transmission processes generating both series are likely to be different, raising additional
questions regarding the significance and extent of the risk of zoonotic transmission of *Giardia* infection between dogs and people.

Conflict of interest statement: The authors declare no conflict of interest.
References:


Table Legends

Table 1: Coefficients of the non-seasonal and seasonal terms, with corresponding standard errors (s.e.) of SARIMA models for human and canine data from 2003-2009.
Figure captions

Figure 1: Time plots of human giardiasis (A) reported to CDC and percent of positive fecal canine tests (B) at Banfield Pet Hospital, by month, United States 2003 -- 2009.

Figure 2A: STL decomposition of human giardiasis by month, United States, 2003 – 2009, into trend, seasonal and remainder components. The y-axis scale represents cases/100,000.

Figure 2B: Plots of average monthly incidence of human giardiasis cases/100,000 (y-axis) from the surveillance raw data (top) and STL seasonal component (bottom), United States, 2003 - 2009.

Figure 3A: STL decomposition of canine giardiasis at Banfield hospitals, 2003-2009, into trend, seasonal and remainder components. The y-axis scale represents positive fecal tests/100 dogs.

Figure 3B: Plots of average monthly prevalence of positive fecal tests (y-axis) among dogs from the raw data (top) and STL seasonal component (bottom).
Figure 1: Time plots of human giardiasis (A) reported to CDC and percent of positive fecal canine tests (B) at Banfield Pet Hospital, by month, United States 2003 -- 2009.
Figure 2A: STL decomposition of human giardiasis by month, United States, 2003 – 2009, into trend, seasonal and remainder components. The y-axis scale represents cases/100,000.
Figure 2B: Plots of average monthly incidence of human giardiasis cases/100,000 (y-axis) from the surveillance raw data (top) and STL seasonal component (bottom), United States, 2003 - 2009.
Figure 3A: STL decomposition of canine giardiasis at Banfield hospitals, 2003-2009, into trend, seasonal and remainder components. The y-axis scale represents positive fecal tests/100 dogs.
Figure 3B: Plots of average monthly prevalence of positive fecal tests (y-axis) among dogs from the raw data (top) and STL seasonal component (bottom).
Table 1: Coefficients of the non-seasonal and seasonal terms, with corresponding standard errors (s.e.) of SARIMA models for human and canine data from 2003-2009.

<table>
<thead>
<tr>
<th></th>
<th>Human series</th>
<th>Canine series</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>SARIMA(0,0,1)x(0,1,1)_{12}</td>
<td>SARIMA(0,1,1)x(0,0,1)_{4}</td>
</tr>
<tr>
<td>Term</td>
<td>Coefficient</td>
<td>s.e.</td>
</tr>
<tr>
<td>Non-seasonal</td>
<td>0.2015</td>
<td>0.124</td>
</tr>
<tr>
<td>Seasonal</td>
<td>-0.6231</td>
<td>0.158</td>
</tr>
</tbody>
</table>