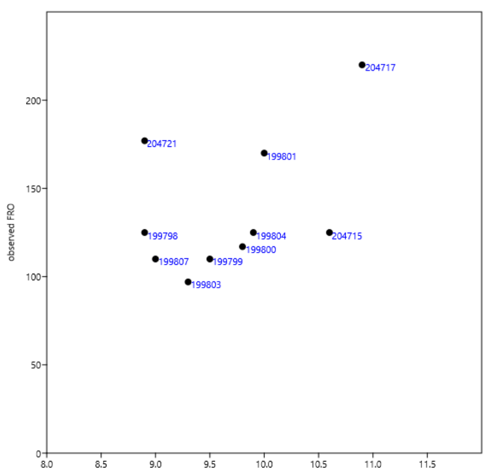
Overall I think this is a convincing study which merits publication. It's a well-organised manuscript with all the essential information readily accessible. The pattern of dog ^{14}C ages is inexplicable without dietary reservoir effects of the magnitude proposed by the authors, and the suggested diet reconstructions would easily account for such dietary reservoir effects in many freshwater systems (i.e. there are many freshwater systems in which modern fish has an apparent ^{14}C age of >400y, as implied by the results and models proposed in this manuscript). I completely agree with the authors' interpretation, and I believe they are right to question the use of uncorrected ^{14}C ages on human and dog samples from Ontario.

I have made a few suggestions that I think could improve the manuscript.

The main problem is one of circular reasoning: our idea of what magnitude of dietary reservoir effect is reasonable in this context does not depend on independent measurements (e.g. of the ^{14}C ages of known-age fish from the same ecosystems) but on the pattern of dog ^{14}C ages, relative to the ^{14}C ages of contemporaneous deer/maize samples. The only independent evidence for the scale of the potential freshwater reservoir effects (FREs) is a rather rubbery correlation (published in 2013) between freshwater alkalinity and modern fish ^{14}C ages in other parts of the world. Were there no fish bones available to date from the same sites in Jefferson County, which would have given a reasonable estimate of the potential FREs in local fish?

Theoretically, we could actually use the data presented in this paper to calculate a local FRE, using the linear-regression approach, as used with paired human bone-grave good dates by Schulting et al. 2014 (cited here). This does not require accurate diet reconstruction, merely a predictable correlation between (one or more of) the stable isotopes and the FROs in dog bones. I am surprised that the authors haven't mentioned this method. Scatter plots of stable isotopes vs FROs (combining data in Tables 2 and 4) don't show a convincing correlation – unsurprising for $\delta^{13}C$, which will be dominated by the C3 vs C4 plant signal, but $\delta^{15}N$ doesn't appear to explain much of the variation in FROs either:



This is at first surprising, given that nearly all the fish samples in figure 2 have higher $\delta^{15}N$s than nearly all the terrestrial species. I can think of at least 3 factors which might obscure a relationship between dog $\delta^{15}N$ and FRO:

- localised differences in fish $\delta^{15}N$, and more importantly, FRE
- differences in $\delta^{15}N$ enrichment factors between dogs (e.g. if the data set includes very young dogs that were not fully weaned)
- inclusion of some samples that were not approximately contemporaneous with other samples from the same site (which is impossible for anyone to judge, given the excavation history of these sites)

The MixSIAR models with 7 food sources seem unnecessarily complex, and rather arbitrary; as the focus here is on radiocarbon, perhaps it is legitimate to combine the fish categories, since (as far as we know) there is no difference in FRE between high, medium and low $\delta^{15}N$ fish. Does the 7-source model predict higher fish consumption than the 3-source model? Why would this be?

I was surprised that there was no mention of C3 plant foods in the MixSIAR models (which isotopically might look like your small herbivores, i.e. your small herbivore estimates might include consumption of C3 plants) and no mention of animal domesticates (e.g. was the turkey domesticated by the 15th century? if so, are the turkey stable isotope data relevant, as they seem to reflect almost pure C3 diets?)

The OxCal bounded-phase models are a bit redundant, to be honest. I guess the authors are trying to make the point that including calibrated uncorrected dates of dog (and by implication human) bones will produce misleadingly early chronologies. This can be done in multiple ways. I am not sure how the Delta_R for dogs was calculated (seems to be the mean ± standard deviation of the individual FROs, whereas the text mentions the weighted mean FRO, which is lower and with a smaller uncertainty), but logically you would either use a different Delta_R for each dog, depending on its diet reconstruction, or a really vague likelihood for Delta_R, such as U(0,500), allowing the model to compile all possible values consistent with the dogs dating to the same phase as the maize and deer samples. Either way, the important thing would be to show the model's posterior estimate of Delta_R, showing how far the dog ^{14}C ages have to be shifted in order to fit a chronology based on the dates of deer and maize samples (e.g. 167±14y in my model). However, I would first use the R_Combine function (in reality the Ward and Wilson (1978) test) to check whether any of the uncalibrated dog dates are statistically consistent with the terrestrial samples from the same sites, just to show that this is never the case, and that the dog dates are therefore significantly older.

Please check that references are listed in alphabetical order in the bibliography!

Some detailed line-by-line comments on the text:

6-7: "Freshwater bodies can sequester ancient carbon…." - wording a bit obscure (I know what you mean, but others might not); what about e.g. "Carbon entering the food web in freshwater systems is often not in full isotopic equilibrium with the atmosphere, giving rise to spuriously old radiocarbon ages in fish and other aquatic organisms"
Similar reaction to line 36 - I think that the wording is too general for readers who don't already understand the mechanisms involved, i.e., that the food-web in freshwater

systems assimilates C via photosynthesis of DIC (dissolved inorganic carbon, formed by primarily dissolution of carbonate minerals in the catchment) and recycling of aged organic C (the division between dissolved organic carbon and particulate organics is rather arbitrary). There is also some influx of fresh organic carbon from terrestrial sources (recent leaves etc) and quite a bit of DIC exchange with atmospheric CO_2, both of which serve to “rejuvenate” the freshwater C.

11-12: (and subsequently) “cal. mid-fifteenth to mid-sixteenth- century” - not sure that “cal.” helps here; 15^{th}-16^{th} century should be sufficiently unambiguous, given the context

73-75: “Having established that fish accounted for 20-40% of the diets of dogs in our sample from these villages, we next determine if radiocarbon dates on dog bones result in FROs. We do this by comparing ^{14}C ages…” – the first sentence seems out of place here, and might go better at lines 53-54. However, it is difficult to follow the 20-40% estimate, given that this is still just the introduction (“having established that…” cannot really appear before the results section). The next sentence (lines 75ff) would need to be reworded slightly, but not much – the FROs in dogs follow from the conclusion that each site was short-lived

89: “Bone fragments were submitted…” - just dog bone, or deer bone as well?

106: “Data for prey (source) animal species were obtained from Booth (2014), Guiry et al. (2016), Morris (2017), Morris et al. (2016) and Pfeiffer et al. (2016).” - how does the date range of this material compare to the dates of your samples? can you be confident that isotope baselines were constant through time? I also think you need to make an explicit argument for the applicability of Ontario fish isotopes in Jefferson Co. NY.

Likewise, in lines 149ff: “.. the isotope values for dogs from Jefferson County plots among those for dogs from southern Ontario…” – are the Ontario dogs from the same period as the Jefferson County dogs?
“… These results suggest use of isotopes from southern Ontario in the Bayesian mixing models for potential prey sources is warranted.” – similarity is unsurprising for deer (isotope values will depend mainly on vegetation), but why should fish values be equivalent? similar geology and climate?

111: “The diversity of presumed food was placed into seven sources in the mixing models.” odd wording - what about “potential food sources were placed in 7 groups in the mixing models”?

116: “fish were split into three separate sources based on statistically significant differences among taxa in $\delta^{15}N$ values,” - unclear how this was done, as their ranges evidently overlap, even if the mean/median $\delta^{15}N$s for these arbitrarily defined groups are significantly different. Why 3 groups, and not 2 or 4? it might be logical to create 3 if they represent different habitats, fishing techniques etc., but no explanation is provided.

124: “trophic enrichment factors (TEF) for $\delta^{13}C$ (+1.1‰ ± 0.2‰) and $\delta^{15}N$ (+3.8‰ ±1.1‰) were derived from Ledogar et al. (2018) and Bocherens et al. (2015).” - do not cite Ledogar et al. here, as they only copy the TEF calculated by Bocherens et al. Do you know of any study (e.g. feeding experiment) that specifically considered TEF for dogs? Do you know whether all the dog samples are from tissues remodeled in adulthood, or could there be a remnant nursing effect in the $\delta^{15}N$ values?

125: “The maize (source) to consumer (dogs) TEF was +5.0‰ ± 0.1‰ for $\delta^{13}C$, and +3.0‰ ± 0.1‰ for $\delta^{15}N$.” – how were these numbers obtained?

127: “models were only accepted when the criteria of the Gelman-Rubin and Geweke tests were met” - what are these tests for? did some variants of the models give poor convergence, and how would you interpret that?

138: “In order to graphically display the results of the modelling we use Phase Start and End boundaries” - odd justification; the use of boundaries imposes a statistical model on the dates (they should be sampled from a uniform distribution beginning and ending at the boundaries), which is a reasonable assumption to make when the samples are in some way associated with each other

139: “and within Phase undated event and KDE plot” - clearly these functions (OxCal Date function and KDE_Plot function) give practically identical results, so only one (at most) is really necessary; I would choose one and explain your choice.

142: “FRO standard deviations were calculated by adding together those of the dates used in the FRO calculations” - no, Keaveney and Reimer followed the standard scientific practice of calculating the uncertainty in the difference between two measurements as the square root of the sum of the squared uncertainties in individual measurements, i.e. $\sigma_{FRO} = \sqrt{\ } (\sigma_{dog}^2+\sigma_{deer}^2)$. Please recalculate your uncertainties following this method.

145: “combined with the OxCal R_Combine(); for FRO calculations” - i.e. they were compared according to Ward and Wilson (1978), and, as in each case the dates were not significantly different at the 5% significance level, their weighted mean was used for FRO calculations. See OxCal manual for W&W reference.

164: “results for individual dogs are presented in Supplement Table 3.” - attached file is called Supplementary File 3 but has the heading Supplemental File 1

168: “Results of radiocarbon assays on dog and contextual deer bone and maize are presented in Table 2” - Table 2 also includes calibrations – which is fine for terrestrial samples, but clearly (considering the purpose of this paper) completely pointless for the dog results. I would omit the calibrated date column

173: “These results clearly indicate the ^{14}C ages of dog bone are older than those of deer bone and maize.” - this seems like an unnecessary mixing of units. To observe FROs, the dog ^{14}C ages can be compared directly to the deer/maize ^{14}C ages, as in Table 4; converting the results to calendar ages first could be confusing, depending on the shape of the calibration curve. I would leave the calibrations to the discussion section, and here focus on the ^{14}C results themselves.

190: "dog bone fraction dead carbon (FDC) of 0.0115 to 0.0252." - not sure that this is necessary (sounds like a reformulation of the FRO values), but if it is, you need to explain how it was calculated. It looks like a simple subtraction of the dog F14C from the deer/maize F14C. The dog's "dead carbon fraction" should be ($1 - F14C_{dog}/F14C_{deer/maize}$), which gives slightly higher FDC numbers.

191: "using a ΔR prior" - technically you have specified a likelihood for ΔR, not a prior; but it's also a circular argument (you used the dog vs deer/maize FROs to choose a restricted likelihood for ΔR, so the calibrated dog dates *must* now agree with the calibrated deer/maize dates, by definition); it would be more informative to specify a vague likelihood for ΔR, allowing also for negligible FROs, then look at the posterior distribution of this parameter (which will be determined by the data)

196: "Estimates of mean FDCs for fish consumed by individual dogs were calculated" - this might be the justification for calculating the dog FDCs, i.e. that you can use the dog FDCs scaled according to the %fish in dog diets to estimate the FRO in the fish eaten by these dogs. With my dog FDC formula, treating each dog independently, I get 400-900y FROs for fish (average 500-600y), which sounds reasonable. Most of the scatter in the individual estimates of fish FRO might be attributable to the relatively large uncertainties in the MixSIAR output (although there's also a significant uncertainty in the dog FDC results), rather than real variation in fish FROs.

198: "It is probable, then, that total alkalinity values were >90 mg/L in area fisheries during the time in question." - this is not the most compelling conclusion you could draw! If you plot dog FDC vs estimated fish intake from the MixSIAR models, the values should be highly correlated, if the fish consumed by all dogs had the same FDC/FRE/FRO. There's only a modest positive correlation, however, which means either that the MixSIAR estimates are noisier than we would like, or that the fish aren't all from the same sources (see previous comment)

213: "the headwaters of the St. Lawrence River, an area with high total alkalinity,.." - is it? compared to what? probably not necessary anyway. There is a risk that readers assume there are no reservoir effects in soft water, which is not true.

254: "Archaeologists pursuing the canine surrogacy approach to assess human diets through isotopic analysis should take the possibility of FROs into account prior to obtaining radiocarbon dates…" - I would also be more positive, by saying that you can also use ^{14}C/FRO as a dietary tracer in some cases (dating articulated dog bones accompanied by fully terrestrial samples, if you know that local fish has a significant FRO)