A Linked Mixture Model of Coronary Atherosclerosis

A Preprint

Bret Beheim
Department of Human Behavior, Ecology and Culture
Max Planck Institute for Evolutionary Anthropology
Leipzig, Germany
bret_beheim@eva.mpg.de

July 11, 2021

Abstract

Background & Objectives: Characterizing the progression of coronary atherosclerosis is a critical public health goal. The most common quantitative summary, the CAC score, is modelled by a variety of statistical methods, both as a predictor of coronary events and as an outcome of behavioral and population-specific risk factors. Little attempt has been made, however, to ground these statistical models in the underlying physiology of arterial aging, which would allow us to describe the onset and growth of CAC over a patient's life.

Methods: Using a stochastic growth model for arterial plaque accumulation, I identify severe under-estimation in the age of initial onset and rate of progression (doubling time) of CAC growth, and use this generative growth model to motivate new statistical approaches to CAC using mixture models of logistic and log-linear outcomes.

Results: Using 640 growth trajectory simulations (71,680 patients, 277,600 observations), I demonstrate substantial improvements in predictive accuracy of CAC onset and progression with linked hurdle-lognormal mixture distributions, compared with standard ln(CAC+1) or ln(CAC) models.

Conclusions: Using models that can account for patient-specific onset and progression rates, accurate descriptions of CAC trajectories can be made even in cross-sectional (single scan per patient) designs, with substantial clinical and epidemiological utility.

Keywords Agatston Scores · Coronary Calcium · Bayesian Methods · Generative Inference

1 Background & Objectives

The coronary calcium (CAC) score was first proposed by Agatson, *et al.* [1] to quantify the amount of calcified plaque in the coronary arteries, and remains the standard metric for assessing arterial aging [2]. Three decades of research using large cohorts of patients, in both cross-sectional and longitudinal designs, has shown a patient's CAC score to be highly predictive of subsequent major coronary events [3–7], even among patients with few additional risk factors [8, 9]. Many studies have found CAC varies systematically by nationality [10, 11] and ethnicity [12–15], with some populuations showing very little CAC even into advanced age [16], leading to debates over whether atherosclerosis itself is a consequence of sedentary, industrialized lifestyles [17, 18].

Yet, despite its centrality in the study of atherosclerosis, the CAC score's significance in clinical practice is ambiguous [19, 20], and it lacks a consistent statistical model from study to study. A given analysis may employ ordered categorical [7, 21–23], linear [24, 25], log-linear [20, 26–28], quantile [29, 30] and proportional hazards [31–33] regression models, using CAC both as an outcome [5, 16, 34] or a predictor [32, 35]. With some exceptions, e.g. [25, 29, 36–39], the CAC score is almost always modeled as a cross-sectional value that varies between patients, rather than as the result of a growth process within an individual over the course of their lives.

To better connect the CAC score to the dynamic physiology of arterial aging, I propose a simple growth model for the onset and progression of calcified arterial plaques. Using high-throughput computational simulations (640 studies, 71,680 patients, 277,600 observations), this model can identify structural mis-specifications in many of the common log-linear models of CAC, and indicate specific statistical modifications which dramatically improve the predictive accuracy of CAC modeling across the life course.

2 Methods

2.1 A Generative Model of CAC Growth

Although CAC is modeled in the current literature by a large variety of methods, we can synthesize two key empirical details from which to build a generative growth model of CAC. First, calcified plaques typically appear in the coronary artery in middle age, with substantial variability between patients. Before this age of onset, t_0, patients given a computerized tomography scan would have a CAC score of 0. After the age of onset, an individual's CAC steadily accumulates at one or more plaque sites in the coronary artery, increasing at a non-linear pace set in part by variation in genetics, behavior, or risk exposures [8, 25, 36, 40]. Given this, let us assume that the rate of growth in CAC at a given moment t after the age of onset t_0 is proportional to both instantaneous risk factors and the current amount of coronary calcium, that is,

$$\frac{dy}{dx} = y(t)\, k(t), \tag{1}$$

for CAC score $y(t)$ and positive growth function $k(t)$. The CAC score y is thus given by the integral of all growth velocities since the age of onset t_0,

$$y(t) = \int_{t_0}^{t} y(x)\, k(x)\, dx \tag{2}$$

As we observe the score y change with time, we could estimate the shape of the rate function $k(t)$. However, since we ordinarily do not have more than one or a few CAC observations per person, we must instead appeal to the *average* CAC growth rate between t_0 and t, $k = (t - t_0)^{-1} \int_{t_0}^{t} k(x)\, dx$, a constant. We can then solve Equation 1 as

$$\ln y(t) = k\,(t - t_0) \quad \text{for any} \quad t \geq t_0. \tag{3}$$

Since rate k describes an exponential growth process, it is useful to re-express k as a *doubling time*. That is, we can describe the time $y + d$ at which the outcome score is doubled from a current amount y at time t as $\ln 2y = k\,(t + d - t_0)$, which is true if $d = (\ln 2)/k$. An example atherosclerosis trajectory for a single patient is given in Figure 1.

Allowing for a small amount of measurement error, we can estimate the average k for a set of observations by log-linear regression of CAC score y_j on patient age t_j using the above model. That is,

$$\begin{aligned} y_j &\sim \text{LogNormal}(\mu, \sigma) \\ \mu_j &= k\,(t_j - t_0), \end{aligned} \tag{4}$$

where μ_j represents the average CAC score at age t_j, which varies due to measurement error and other unobserved factors with standard deviation σ. Given a dataset recording a single individual's CAC trajectory over time, and the log-linear regression model $\mu_j = a + bt_j$, we could estimate the above model parameters as $t_0 = (-a/b)$, $k = b$ and $d = \ln(2)/b$.

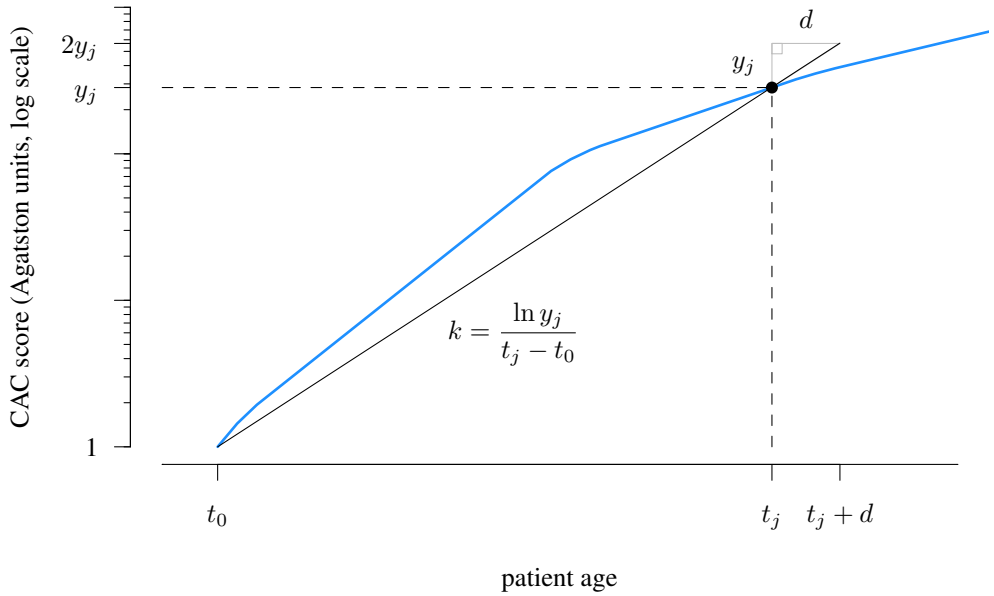

Figure 1: A single patient's coronary atherosclerosis progression over time, as measured by CAC score (blue). Onset at age t_0 is defined as the initial appearance of calcified plaque (CAC > 0), which is equivalent to a CAC score of 1 in most studies. Before t_0 the patient's CAC is 0. All CAC measurements after t_0 are given by Equation 2. In this growth trajectory, the patient's instantaneous CAC growth rate $k(t)$ intermittently decreases. Although neither the location of t_0 nor the rate function $k(t)$ are directly observable, average growth parameters can be estimated by log-linear regression Equation 4 (black line) for the location of t_0, an average k and doubling time $d = \ln(2)/k$.

2.2 Structural Under-estimation of Disease Onset and Progression

Studying CAC on the logarithmic scale was originally proposed by Agatston, *et al.* [1] and remains the most common statistical approach. However, the growth model described in Equation 4 indicates the existence of two major estimation biases with such models (Figure 2). Principally, the existence of distinct CAC age trajectories across patients is not a structural component of log-linear regression, which treats each patient-observation as following a single population average log-linear path from a single age of onset, t_0. As a consequence, standard regression models of CAC systematically under-estimate the age of atherosclerotic onset (Figure 2, left) and progression (Figure 2, right) in direct proportion to between-patient variability. Under realistic growth conditions, an inter-patient variatibility of only ± 3 years age of onset t_0 can induce estimation biases as severe as five decades.

An additional bias is introduced by studies that alter the scores of patients with no atherosclerosis (CAC = 0). Since 0 is undefined on the logarithmic scale, many studies first re-assign these patients a CAC score of 1 [5, 11, 15, 20, 27, 28, 34, 35, 41–43], indicating these patients have already begun atherosclerosis. The common justification given for the log-transform is to normalize the distribution of CAC [1, 11, 15, 32]. Yet depending on the cohort, a majority of patients may have a CAC score of 0. Consequently, such logarithmic outcome distributions are extremely zero-inflated, e.g. Figure 1 in [11]. By estimating the average growth rate k from a population in which some are growing CAC at rate k and others are not growing CAC at all, log-linear regression models introduce additional biases of about -5.9 years on t_0 (Figure 2, left, red) and an additional +1.0 year of CAC doubling time (Figure 2, right, red), versus studies that exclude non-atherosclerotic patients before performing the regression [14, 16, 25, 26, 44, 45].

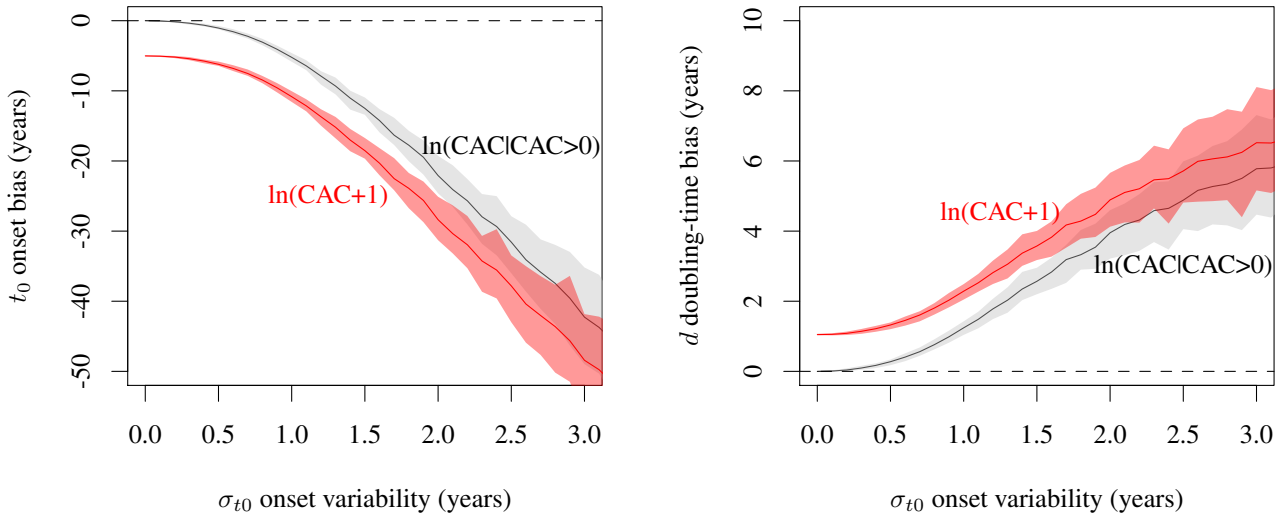

Figure 2: The relationship between inter-individual variance in t_0 and the bias in linear regression estimate of the age of onset (left) and the doubling time of CAC (right) for a sample of single observations from 5,000 patients. Controlling the standard deviation in individual t_0 variability, σ_{t0}, 300 cross-sectional datasets are randomly drawn from 5,000 patient growth trajectories. For each dataset, regression estimates are calculated in a log-linear regression of ln(CAC + 1) (red) or ln(CAC | CAC > 0) (black) on patient age, and the true onset age t_0 in the simulation is subtracted from onset estimates of $t_0 = -a/b$. Average estimates are shown with 89% confidence regions. The ln(CAC + 1) estimates are, on average, -5.9 years more biased than ln(CAC | CAC > 0) on t_0 and +1.0 year more biased in d over the observed range of σ_{t0}. Reproducible R code for this simulation and figure is available in the Supplementary Materials.

2.3 Alternate Statistical Specifications

To prevent inter-individual variation from manifesting as bias in statistical analyses, we can modify the log-linear model in Equation 4 with explicit CAC growth trajectories for each patient i by

$$\begin{aligned} y_j &\sim \text{LogNormal}(\mu_j, \sigma) \\ \mu_j &= k_i\,(t_j - t_{0i}), \\ t_{0i} &\sim \text{Normal}(t_0, \sigma_{t0}), \\ k_i &\sim \text{Normal}(k, \sigma_k). \end{aligned} \tag{5}$$

That is, the time of initial onset for each individual i, t_{0i}, is assumed to vary around some population-average age t_0 by standard deviation σ_{t0}, and similarly for individual-specific average growth trajectory k_i.

Rather than exclude patients who have no CAC, another approach is to employ generalized linear regression to model the probability CAC > 0 [3, 4, 7–10, 31, 39, 46–51]. If we define the log-odds that a patient's CAC score at observation j is non-zero as $\theta_j = \text{logit}(\Pr(y_j > 0))$, then for any individual i observed at age t_j,

$$\theta_j = \frac{t_j - \ell_i}{v} \tag{6}$$

$$\ell_i \sim \text{Normal}(\ell, \sigma_\ell)$$

The probability CAC is observed to be nonzero rises sinusoidally from 0 to 1 over the life course, reaching 50% probability at the inflection point age of ℓ_i for patient i. As above, we can expect ℓ_i to vary by individual according to parameters v and σ_ℓ, but also systematically by risk factors shared across individuals which can be incorporated into the model by logistic regression. Unlike log-linear models above, this approach can utilize all observations, including those for which $y_j = 0$, at the cost of being unable to distinguish non-zero CAC magnitudes.

Models of CAC > 0 can be combined with models of CAC magnitudes in so-called mixture models [52], first introduced in the literature by Kaplan, *et al.* [16] with the use of a zero-inflated Negative Binomial model. Such models incorporate two distinct probability distributions to describe both zero outcomes and non-zero magnitudes simultaneously. We here employ the related concept of a hurdle model, which allows an additional link between these two distributions. Specifically, in modeling the probability of CAC > 0 over the life course, we are simultaneously describing the age of onset as a random quantity. If we re-interpret the inverse-logit link as a cumulative distribution function, onset age must follow a Logistic probability distribution with mean ℓ_i and variance v. Therefore, both the hurdle and log-normal components of this mixture are estimating a common parameter in the individual onset age, t_{0i}, and we can link these stochastic nodes together via

$$t_{0i} \sim \text{Logistic}(\ell_i, v) \tag{7}$$

That is, the initial start times in the log-linear model are random quantities that follow the parameters of the hurdle model, such that the hurdle model's estimates of ℓ_i and v will pass forwards to add additional resolution on t_{0i} in the log-linear model, and the log-linear estimation of t_{0i} will help inform the locations of ℓ_i and b in the hurdle component (Figure 3). Taken together, this fully defines the linked hurdle-lognormal mixture model.[1]

2.4 Ranking Statistical Models by Simulation

In order to characterize gross performance differences between the above methods of studying CAC growth, I simulated 64 populations experiencing a variety of atherosclerotic growth scenarios. Growth trajectories from each population were sampled in 10 datasets, varying in the number of patients (from 1 to 500 patients), the number of observations per patient (in longitudinal or cross-sectional designs), the impact of risk factors of age and sex on initial onset and progression, and the mismatch between population properties and prior information. All simulations were performed in R version 3.6.3 and models were fit with Stan by the `cmdstanr` library version 0.3.0.9. A full list of the permutations is given in Table 2. We can summarize patient data for N = 71,680 patients seen over all 640 studies in Table 1. Each of the 14 models was fit to each of the 640 datasets, statistical estimates were calculated and extracted, and, given the true values for each individual trajectory were known, the bias of each parameter estimate was computed.

3 Results

Model battery performance is summarized in Table 3. In both cross-sectional and longitudinal designs, estimation of onset ages (population average t_0 and patient-specific t_{0i}) is least-biased in models of the binary outcome CAC > 0, including the linked hurdle log-normal and logistic regression models. The (un-linked) hurdle-lognormal and log-linear regression with patient-specific trajectories are consistently biased by nearly two decades across studies, while log-linear regression designs without individual trajectories are biased by at least three decades in estimating onset ages.

Sex specific risk-factor stratifications have small contrast coefficient biases for all models in longitudinal designs, with the largest bias produced by ln(CAC | CAC > 0) models at 3.2 (SE: 1.4) years. In cross-sectional designs, however, bias

[1] A complete list of variables is given in Table 4.

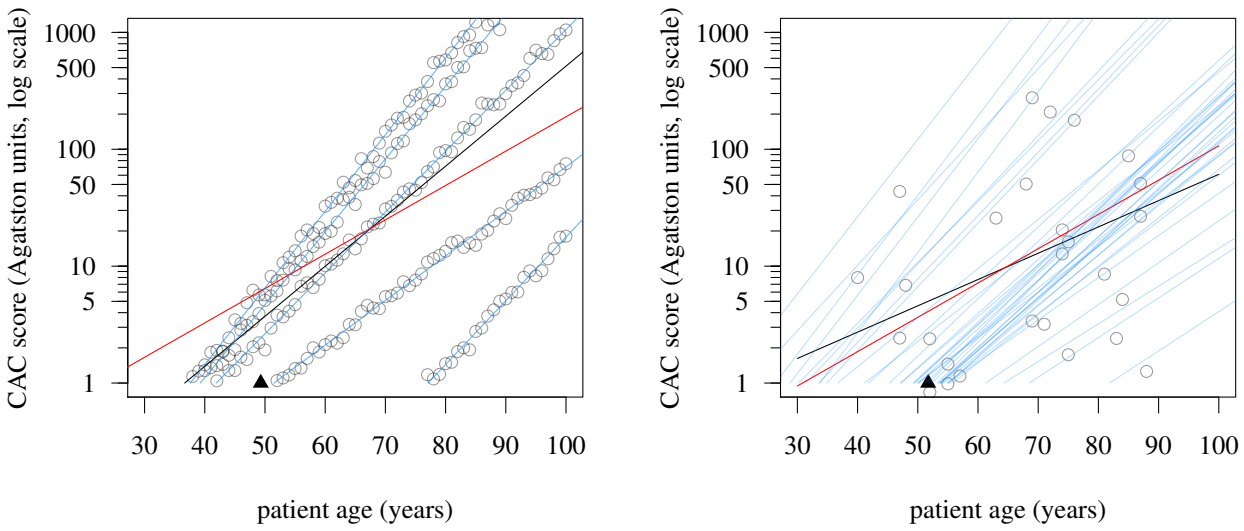

Figure 3: Observed CAC trajectories for 5 longitudinal study patients (left) and 100 cross-sectional patients with single observations each (right), with average counterfactual predictions from fitted ln(CAC) (black), ln(CAC + 1) (red) and linked hurdle-lognormal (blue) models. In each figure, true population average t_0 are indicated by black triangles.

Table 1: Characteristics of patients across 640 simulated CAC studies. Means (standard deviations) are given for t_0 and d. Here longitudinal studies involve 100 annual observations over the entire course of each patient's life (all individuals live to age 100), while cross-sectional studies simulate realistic CAC datasets by randomly sampling a single observation from each patient's adult measurements.

	Women	Men	Total
Number of Patients	36,089	35,591	71,680
Longitudinal	1,043	1,037	2,080
Cross-sectional	35,046	34,554	69,600
Number of Observations	139,346	138,254	277,600
Longitudinal	104,300	103,700	208,000
Cross-sectional	35,046	34,554	69,600
Onset t_0 (years)	48.9 (23.9)	50.2 (23.2)	49.5 (23.6)
Doubling time d (years)	2.7 (1.3)	2.9 (1.4)	2.8 (1.3)
CAC category (%)			
0	51.4	51.8	51.6
1 to 9	9.8	10.1	9.9
10 to 99	8.7	8.7	8.7
100 to 399	4.5	4.4	4.5
400 to 999	2.7	2.7	2.7
$\geq$ 1000	22.7	22.1	22.5

Table 2: Conditions across 640 simulated CAC studies. Here "longitudinal" studies involve 100 annual observations over the entire course of each patient's life (all individuals live to age 100), while "cross-sectional" studies simulate realistic CAC datasets by randomly sampling a single observation from each patient's adult measurements. Sex differences represent a common risk factor across studies and, if present in the population, may stratify the onset and progression by men versus women by a standard deviation of 10 years on t_0 and 112 years on d. Individual variation here is described by model parameter σ_{t0}. A "regularizing" prior reflects realistic ranges for each parameter prior to updating in Stan [52].

	Longitudinal	Cross-sectional
Total Studies	240	400
Inter-Individual Variation		
No Variation	90	0
$\pm$ 3 years on t_0	150	250
Sex Effect		
No sex difference in t_0	60	100
Sex difference in t_0	180	300
Sex difference in t_0, k	120	200
Prior Accuracy		
Regularizing Priors	30	50
Strong Priors	30	50
Perfect Priors	180	300
Study Size		
N = 1 Patient	80	0
N = 5 Patients	80	0
N = 20 Patients	80	80
N = 50 Patients	0	80
N = 100 Patients	0	80
N = 200 Patients	0	80
N = 500 Patients	0	80

is relatively higher for logistic regression models, but especially for the standard log-linear regressions at -9.2 (SE: 4.5) and -8.0 (SE: 27.5) years. The linked hurdle log-normal yields on average the least-biased estimates of a sex effect on age of onset, at -0.3 (SE: 0.7) years.

Estimates of the speed of atherosclerotic progression, measured as doubling times, are possible in log-linear models of CAC magnitudes. Estimation accuracy is high in longitudinal studies, but in cross-sectional studies, bias is only lower for models that incorporate patient-specific individual trajectories. Models without these estimate slower rates of CAC growth than the true rates, with a doubling-time bias of 7.5 (SE: 1.4) years for ln(CAC + 1) and 10.9 (SE: 4.8) years for ln(CAC | CAC > 0).

4 Discussion

Generally speaking, logistic and mixture models out-perform log-linear models in accurately estimating population and individual-specific ages of CAC onset, and models that explicitly account for individual patient trajectories out-perform those that do not. These advantages are most pronounced in cross-sectional study designs, which constitute the vast majority of CAC datasets.

Because logistic models are able to accurately utilize the full range of patient observations, including the period before CAC growth begins, they are able to more-accurately characterize the onset of atherosclerosis compared to log-linear models, which must either exclude patients observed with zero CAC or incorrectly alter these scores to accomodate the logarithmic scale, producing large estimation errors.

Table 3: Estimation bias (mean and standard error) over all 640 studies.

model	design	t_0 bias	t_0 sex bias	d bias
logit(Pr(CAC > 0)) + RE	cross-sectional	-6.6 (3.7)	1.5 (3.7)	-
logit(Pr(CAC > 0))	cross-sectional	-7.8 (4.2)	2.1 (4.2)	-
linked hurdle-lognormal	cross-sectional	-8.2 (0.7)	-0.3 (0.7)	1.0 (1.2)
ln(CAC | CAC > 0) + RE	cross-sectional	-20.8 (1.1)	0.5 (1.1)	0.3 (0.1)
hurdle-lognormal	cross-sectional	-23.5 (3.7)	0.7 (3.7)	0.3 (0.1)
ln(CAC + 1)	cross-sectional	-28.8 (4.5)	-9.2 (4.5)	7.5 (1.4)
ln(CAC | CAC > 0)	cross-sectional	-49.9 (27.5)	-8.0 (27.5)	10.9 (4.8)
logit(Pr(CAC > 0))	longitudinal	-0.5 (0.7)	0.7 (0.7)	-
linked hurdle-lognormal	longitudinal	-0.9 (0.4)	0.5 (0.4)	-0.1 (0.0)
logit(Pr(CAC > 0)) + RE	longitudinal	-1.2 (0.4)	-0.4 (0.4)	-
hurdle-lognormal	longitudinal	-2.1 (0.6)	-1.2 (0.6)	-0.1 (0.0)
ln(CAC | CAC > 0) + RE	longitudinal	-2.3 (0.7)	-0.3 (0.7)	-0.1 (0.0)
ln(CAC | CAC > 0)	longitudinal	-22.2 (1.4)	3.2 (1.4)	0.9 (0.1)
ln(CAC + 1)	longitudinal	-31.0 (0.9)	1.0 (0.9)	3.0 (0.3)

Log-linear models are also observed to produce a large biases in estimating the impact of risk factors across these simulation studies. Such biases in modeling the age of onset can potentially mask the importance of predictors with measurable statistically significance, or even produce spurious correlations.

Although the innovations in model design described above show marked improvements in estimation accuracy, they come with increased costs, in terms of model complexity and computation time, and successful posterior updating of the full linked-hurdle log-normal may not be feasible for very large samples of patients. However, Table 3 indicates that estimation performance can be improved for specific parameters without the use of the full linked hurdle-lognormal model: for parameter d, the ln(CAC | CAC > 0) + RE model is an improvement, and for t_0, the logistic regression models.

More generally, although the simulation studies presented here cover a broad range of realistic epidemiological scenarios, they are strictly limited to the context of the generative model described by Equations 1 and 2. The true process of CAC onset and progression is necessarily more complex, and is only part of the larger process of coronary atherosclerosis, which can be seen as beginning with arterial lesions and soft plaques, before the appearance of any calcification. Incorporating more physiological details, such as the process of initial calcification of soft plaque, or the site-specific formation of individual calcified plaques, can potentially lead to further modeling innovations, and consequently greater improvements in estimation accuracy.

5 Conclusion

Review of the current literature shows subtantial variability in the methods of studying the CAC score, with multiple statistical designs employed even within the same study. Both because of this awkwardness, and because the CAC score aggregates over distributions of calcified plaque, several studies have suggested alternative metrics to studying atherosclerosis progression [19, 20]. However, the above computational meta-analysis shows the current statistical approaches to CAC modeling have substantial structural biases, tending to represent atherosclerotic growth as beginning earlier and progressing slower than it really does. Modifications to these models that recognize patient-specific growth trajectories are substantially more accurate in realistic simulation studies with both longitudinal or cross-sectional designs, and so we can expect such models to similarly out-perform log-linear models in the analysis of real patient data.

Table 4: Glossary of variables used.

Variable	Description
$k(t)$	instantaneous growth rate in CAC at age t
$y(t)$	true patient CAC score at age t
y_j	observed CAC score at observation j
t_j	patient age at observation j
t_0	population average age of onset
t_{0i}	age of onset for patient i
k	population average growth rate in CAC
d	population average dounling time in CAC
k_i	average growth rate in CAC for patient i
d_i	average doubling time in CAC for patient i
a	log-linear regression intercept
b	log-linear regression slope
v	logistic slope on logit(Pr(CAC > 0))
ℓ	expected population average age of onset
ℓ_i	logistic inflection point for logit(Pr(CAC > 0))
θ_j	log-odds of CAC > 0 for patient i in measurement j
σ	residual standard deviation on each CAC measurement y_j
σ_{t0}	standard deviation on t_{0i}
σ_k	standard deviation on k_i

6 Acknowledgments

Hillard Kaplan, Michael Gurven, Gregory Thomas, Michael Miyamoto, Richard McElreath, the Department of Human Behavior, Ecology and Culture at MPI-EVA provided valuable feedback on earlier versions of this analysis.

7 Supplementary Materials

All statistical models and simulation data and software is available at https://github.com/babeheim/linked-calcium-growth under the CC BY-NC-SA 4.0 license.

References

[1] Arthur S. Agatston et al. "Quantification of coronary artery calcium using ultrafast computed tomography". In: *Journal of the American College of Cardiology* 15.4 (1990). Publisher: American College of Cardiology Foundation, pp. 827–832. DOI: 10.1016/0735-1097(90)90282-T.

[2] Jarett D Berry et al. "Lifetime Risks of Cardiovascular Disease". In: *New England Journal of Medicine* (2012), p. 9. DOI: 10.1056/NEJMoa1012848.

[3] R Vliegenthart. "Coronary calcification detected by electron-beam computed tomography and myocardial infarction. The Rotterdam Coronary Calcification Study". In: *European Heart Journal* 23.20 (2002), pp. 1596–1603. DOI: 10.1053/euhj.2002.3240.

[4] George T. Kondos et al. "Electron-Beam Tomography Coronary Artery Calcium and Cardiac Events: A 37-Month Follow-Up of 5635 Initially Asymptomatic Low- to Intermediate-Risk Adults". In: *Circulation* 107.20 (2003), pp. 2571–2576. DOI: 10.1161/01.CIR.0000068341.61180.55.

[5] Yadon Arad et al. "Coronary calcification, coronary disease risk factors, C-reactive protein, and atherosclerotic cardiovascular disease events: the St. Francis Heart Study". In: *Journal of the American College of Cardiology* 46.1 (2005), pp. 158–165. DOI: 10.1016/j.jacc.2005.02.088.

[6] Michael J. LaMonte et al. "Coronary Artery Calcium Score and Coronary Heart Disease Events in a Large Cohort of Asymptomatic Men and Women". In: *American Journal of Epidemiology* 162.5 (2005), pp. 421–429. DOI: 10.1093/aje/kwi228.

[7] Matthew J. Budoff et al. "Long-Term Prognosis Associated With Coronary Calcification". In: *Journal of the American College of Cardiology* 49.18 (2007), pp. 1860–1870. DOI: 10.1016/j.jacc.2006.10.079.

[8] Allen J. Taylor et al. "Coronary Calcium Independently Predicts Incident Premature Coronary Heart Disease Over Measured Cardiovascular Risk Factors". In: *Journal of the American College of Cardiology* 46.5 (2005), pp. 807–814. DOI: 10.1016/j.jacc.2005.05.049.

[9] Susan G. Lakoski. "Coronary Artery Calcium Scores and Risk for Cardiovascular Events in Women Classified as "Low Risk" Based on Framingham Risk Score: The Multi-Ethnic Study of Atherosclerosis (MESA)". In: *Archives of Internal Medicine* 167.22 (2007), p. 2437. DOI: 10.1001/archinte.167.22.2437.

[10] Raul D. Santos et al. "Difference in atherosclerosis burden in different nations and continents assessed by coronary artery calcium". In: *Atherosclerosis* 187.2 (2006), pp. 378–384. DOI: 10.1016/j.atherosclerosis.2005.09.017.

[11] Axel Schmermund et al. "Comparison of subclinical coronary atherosclerosis and risk factors in unselected populations in Germany and US-America". In: *Atherosclerosis* 195.1 (2007), e207–e216. DOI: 10.1016/j.atherosclerosis.2007.04.009.

[12] Julie Anne Hoff et al. "Age and gender distributions of coronary artery calcium detected by electron beam tomography in 35,246 adults". In: *The American Journal of Cardiology* 87.12 (2001), pp. 1335–1339. DOI: 10.1016/S0002-9149(01)01548-X.

[13] D. E. Bild. "Multi-Ethnic Study of Atherosclerosis: Objectives and Design". In: *American Journal of Epidemiology* 156.9 (2002), pp. 871–881. DOI: 10.1093/aje/kwf113.

[14] Diane E. Bild et al. "Ethnic Differences in Coronary Calcification: The Multi-Ethnic Study of Atherosclerosis (MESA)". In: *Circulation* 111.10 (2005), pp. 1313–1320. DOI: 10.1161/01.CIR.0000157730.94423.4B.

[15] Matthew J. Budoff et al. "Ethnic differences of the presence and severity of coronary atherosclerosis". In: *Atherosclerosis* 187.2 (2006), pp. 343–350. DOI: 10.1016/j.atherosclerosis.2005.09.013.

[16] Hillard Kaplan et al. "Coronary atherosclerosis in indigenous South American Tsimane: a cross-sectional cohort study". In: *The Lancet* 389.10080 (2017), pp. 1730–1739. DOI: 10.1016/S0140-6736(17)30752-3.

[17] Randall C Thompson et al. "Atherosclerosis across 4000 years of human history: the Horus study of four ancient populations". In: *The Lancet* 381.9873 (2013), pp. 1211–1222. DOI: 10.1016/S0140-6736(13)60598-X.

[18] Michael D. Gurven and Daniel E. Lieberman. "WEIRD bodies: mismatch, medicine and missing diversity". In: *Evolution and Human Behavior* 41.5 (2020), pp. 330–340. DOI: 10.1016/j.evolhumbehav.2020.04.001.

[19] Daniel S. Berman, Yoav Arnson, and Alan Rozanski. "Coronary Artery Calcium Scanning". In: *JACC: Cardiovascular Imaging* 9.12 (2016). Publisher: American College of Cardiology Foundation, pp. 1417–1419. DOI: 10.1016/j.jcmg.2016.05.020.

[20] Blaha Michael J. et al. "Improving the CAC Score by Addition of Regional Measures of Calcium Distribution". In: *JACC: Cardiovascular Imaging* 9.12 (2016). Publisher: American College of Cardiology Foundation, pp. 1407–1416. DOI: 10.1016/j.jcmg.2016.03.001.

[21] G Koulaouzidis et al. "Aggressive and diffuse coronary calcification in South Asian angina patients compared to Caucasians with similar risk factors". In: *International Journal of Cardiology* (2013), p. 5. DOI: 10.1016/j.ijcard.2012.05.102.

[22] M. G. Silverman et al. "Impact of coronary artery calcium on coronary heart disease events in individuals at the extremes of traditional risk factor burden: the Multi-Ethnic Study of Atherosclerosis". In: *European Heart Journal* 35.33 (2014), pp. 2232–2241. DOI: 10.1093/eurheartj/eht508.

[23] Akira Fujiyoshi et al. "Cross-sectional comparison of coronary artery calcium scores between Caucasian men in the United States and Japanese men in Japan: the multi-ethnic study of atherosclerosis and the Shiga epidemiological study of subclinical atherosclerosis". In: *American Journal of Epidemiology* 180.6 (2014), pp. 590–598. DOI: 10.1093/aje/kwu169.

[24] G. Koulaouzidis, P.J. Jenkins, and T. McArthur. "Comparison of coronary calcification among South Asians and Caucasians in the UK". In: *International Journal of Cardiology* 168.2 (2013), pp. 1647–1648. DOI: 10.1016/j.ijcard.2013.02.005.

[25] Matthew J. Budoff et al. "Progression of Coronary Calcium and Incident Coronary Heart Disease Events". In: *Journal of the American College of Cardiology* 61.12 (2013), pp. 1231–1239. DOI: 10.1016/j.jacc.2012.12.035.

[26] Michael H. Criqui et al. "Calcium Density of Coronary Artery Plaque and Risk of Incident Cardiovascular Events". In: *JAMA* 311.3 (2014), p. 271. DOI: 10.1001/jama.2013.282535.

[27] Robyn L. McClelland et al. "10-Year Coronary Heart Disease Risk Prediction Using Coronary Artery Calcium and Traditional Risk Factors". In: *Journal of the American College of Cardiology* 66.15 (2015), pp. 1643–1653. DOI: 10.1016/j.jacc.2015.08.035.

[28] Jan M Hughes-Austin et al. "Relationship of coronary calcium on standard chest CT scans with mortality". In: *JACC: Cardiovascular Imaging* 9.2 (2016), pp. 152–159. DOI: 10.1016/j.jcmg.2015.06.030.

[29] Haitham M. Ahmed et al. "Low-Risk Lifestyle, Coronary Calcium, Cardiovascular Events, and Mortality: Results From MESA". In: *American Journal of Epidemiology* 178.1 (2013), pp. 12–21. DOI: 10.1093/aje/kws453.

[30] Yuki Ohmoto-Sekine et al. "Prevalence and distribution of coronary calcium in asymptomatic Japanese subjects in lung cancer screening computed tomography". In: *Journal of Cardiology* 67.5 (2016), pp. 449–454. DOI: 10.1016/j.jjcc.2015.06.010.

[31] Khurram Nasir et al. "Ethnic Differences in the Prognostic Value of Coronary Artery Calcification for All-Cause Mortality". In: *Journal of the American College of Cardiology* 50.10 (2007), pp. 953–960. DOI: 10.1016/j.jacc.2007.03.066.

[32] Aaron R. Folsom. "Coronary Artery Calcification Compared With Carotid Intima-Media Thickness in the Prediction of Cardiovascular Disease IncidenceThe Multi-Ethnic Study of Atherosclerosis (MESA)". In: *Archives of Internal Medicine* 168.12 (2008), p. 1333. DOI: 10.1001/archinte.168.12.1333.

[33] Zhi-hui Hou et al. "Prognostic Value of Coronary CT Angiography and Calcium Score for Major Adverse Cardiac Events in Outpatients". In: *JACC: Cardiovascular Imaging* 5.10 (2012), pp. 990–999. DOI: 10.1016/j.jcmg.2012.06.006.

[34] H S. Oei. "Risk factors for coronary calcification in older subjects The Rotterdam Coronary Calcification Study". In: *European Heart Journal* 25.1 (2004), pp. 48–55. DOI: 10.1016/j.ehj.2003.10.008.

[35] Robert Detrano et al. "Coronary Calcium as a Predictor of Coronary Events in Four Racial or Ethnic Groups". In: *New England Journal of Medicine* (2008), p. 10. DOI: 10.1056/NEJMoa072100.

[36] Robyn L. McClelland et al. "Arterial Age as a Function of Coronary Artery Calcium (from the Multi-Ethnic Study of Atherosclerosis [MESA])". In: *The American Journal of Cardiology* 103.1 (2009), pp. 59–63. DOI: 10.1016/j.amjcard.2008.08.031.

[37] Andrew P. DeFilippis et al. "The Association of Framingham and Reynolds Risk Scores With Incidence and Progression of Coronary Artery Calcification in MESA (Multi-Ethnic Study of Atherosclerosis)". In: *Journal of the American College of Cardiology* 58.20 (2011), pp. 2076–2083. DOI: 10.1016/j.jacc.2011.08.022.

[38] William Arguelles et al. "Relationship of change in traditional cardiometabolic risk factors to change in coronary artery calcification among individuals with detectable subclinical atherosclerosis: The multi-ethnic study of atherosclerosis". In: *International Journal of Cardiology* 174.1 (2014), pp. 51–56. DOI: 10.1016/j.ijcard.2014.03.137.

[39] Seamus P. Whelton et al. "Predictors of Long-Term Healthy Arterial Aging". In: *JACC: Cardiovascular Imaging* 8.12 (2015), pp. 1393–1400. DOI: 10.1016/j.jcmg.2015.06.019.

[40] Matthew J. Budoff et al. "Coronary Calcium Predicts Events Better With Absolute Calcium Scores Than Age-Sex-Race/Ethnicity Percentiles". In: *Journal of the American College of Cardiology* 53.4 (2009), pp. 345–352. DOI: 10.1016/j.jacc.2008.07.072.

[41] Stefan Möhlenkamp et al. "Quantification of Coronary Atherosclerosis and Inflammation to Predict Coronary Events and All-Cause Mortality". In: *Journal of the American College of Cardiology* 57.13 (2011), pp. 1455–1464. DOI: 10.1016/j.jacc.2010.10.043.

[42] Nathan D. Wong et al. "Abdominal aortic calcium and multi-site atherosclerosis: The Multiethnic Study of Atherosclerosis". In: *Atherosclerosis* 214.2 (2011), pp. 436–441. DOI: 10.1016/j.atherosclerosis.2010.09.011.

[43] Maarten J.G. Leening et al. "Coronary Calcification and the Risk of Heart Failure in the Elderly". In: *JACC: Cardiovascular Imaging* 5.9 (2012), pp. 874–880. DOI: 10.1016/j.jcmg.2012.03.016.

[44] Robyn L. McClelland et al. "Distribution of Coronary Artery Calcium by Race, Gender, and Age: Results from the Multi-Ethnic Study of Atherosclerosis (MESA)". In: *Circulation* 113.1 (2006), pp. 30–37. DOI: 10.1161/CIRCULATIONAHA.105.580696.

[45] Lawrence F. Bielak et al. "Differences in prevalence and severity of coronary artery calcification between two non-Hispanic white populations with diverse lifestyles". In: *Atherosclerosis* 196.2 (2008), pp. 888–895. DOI: 10.1016/j.atherosclerosis.2007.01.041.

[46] Rozemarijn Vliegenthart et al. "Coronary Calcification Improves Cardiovascular Risk Prediction in the Elderly". In: *Circulation* 112.4 (2005), pp. 572–577. DOI: 10.1161/CIRCULATIONAHA.104.488916.

[47] Axel Schmermund et al. "Population-based assessment of subclinical coronary atherosclerosis using electron-beam computed tomography". In: *Atherosclerosis* 185.1 (2006), pp. 177–182. DOI: 10.1016/j.atherosclerosis.2005.06.003.

[48] Pamela L. Lutsey et al. "Associations of Acculturation and Socioeconomic Status With Subclinical Cardiovascular Disease in the Multi-Ethnic Study of Atherosclerosis". In: *American Journal of Public Health* 98.11 (2008), pp. 1963–1970. DOI: 10.2105/AJPH.2007.123844.

[49] Tochi M. Okwuosa et al. "Factors Associated With Presence and Extent of Coronary Calcium in Those Predicted to Be at Low Risk According to Framingham Risk Score (from the Multi-Ethnic Study of Atherosclerosis)". In: *The American Journal of Cardiology* 107.6 (2011), pp. 879–885. DOI: 10.1016/j.amjcard.2010.10.072.

[50] Jung-Hee Yu et al. "The Relationship of Body Composition and Coronary Artery Calcification in Apparently Healthy Korean Adults". In: *Endocrinology and Metabolism* 28.1 (2013), p. 33. DOI: 10.3803/EnM.2013.28.1.33.

[51] Victor Aboyans et al. "Lower extremity peripheral artery disease in the absence of traditional risk factors. The Multi-Ethnic Study of Atherosclerosis". In: *Atherosclerosis* 214.1 (2011), pp. 169–173. DOI: 10.1016/j.atherosclerosis.2010.10.011.

[52] Richard McElreath. *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. 2nd ed. Chapman and Hall/CRC, 2020. DOI: 10.1201/9780429029608.