Supplemental Material
Table S1. Model building results.

| Model | Predictors | VIF | AIC   | BIC  | CV-MSE |
|-------|------------|-----|-------|------|--------|
| Step 1: Models including factors of interest and interaction terms |
| 1     | Age, Sex, Weight, Height | All < 10 | -1352.9 | -1347.5 | 0.0282 |
| 2     | Age, Sex, Weight, Height, BMI | BMI, Weight, Height > 10 | -1558.5 | -1553.1 | 0.0240 |
| 3     | Age, Sex, Weight, Height, Weight*Height | All < 10 | -1580.3 | -1574.9 | 0.0239 |
| 4     | Age, Sex, Weight, Height, BMI, weight*height | BMI, Weight, Height, Weight*Height > 10 | -1572.7 | -1567.2 | 0.0238 |
| Step 2: Consideration of higher order terms (added to model 3) |
| 5     | Age, Age², Sex, Weight, Height, weight*height | All < 10 | -1564.3 | -1558.9 | 0.0239 |
| 6     | Age, Sex, Weight, Height, Weight² weight*height | All < 10 | -1636.7 | -1631.2 | 0.0231 |
| 7     | Age, Sex, Weight, Height, Weight² weight*height, Weight³*height | All < 10 | -1645.2 | -1639.8 | 0.0227 |
| 8     | Age, Sex, Weight, Height, Height² weight*height | Height², Height*Weight, Height³*Weight > 10 | -1582.9 | -1577.5 | 0.0239 |
| 9     | Age, Sex, Weight, Height, Height² weight*height, weight³*height² | Height², Height*Weight, Height³*Weight > 10 | -1576.3 | -1570.9 | 0.0239 |
| Step 3: Consideration of additional interaction terms with age (added to model 7) |
| 10    | Age, Sex, Weight, Height, Weight² weight*height, Weight³*height, Sex*Age | Height², Height*Weight, Height³*Weight > 10 | -1633.7 | -1628.3 | 0.0227 |
| 11    | Age, Sex, Weight, Height, Weight² weight*height, Weight³*height, Weight*Age | Weight, Height², Height*Weight, Weight*Age, and Height³*Weight > 10 | -1629.3 | -1623.8 | 0.0228 |
| 12    | Age, Sex, Weight, Height, Weight² weight*height, Weight³*height, Height*Age | Height, Height², Height*Weight, Height³*Weight > 10 | -1650.3 | -1644.8 | 0.0229 |

VIF, variance inflation factor; AIC, Akaike information criterion; BIC, Bayesian information criterion; CV-MSE, cross-validated mean squared error; BMI, body mass index.
Table S2. Final model results (model 7 from Table S1).

| Variable | Estimate  | 95% CI           | P-value |
|----------|-----------|------------------|---------|
| Intercept| 5.7346    | (5.7046, 5.7647) | <0.0001 |
| Age      | 0.006504  | (0.0054, 0.0076) | <0.0001 |
| Sex      |           |                  |         |
| Male     | -0.1152   | (-0.1365, -0.0940) | <0.0001 |
| Female   | Ref       |                  |         |
| Weight   | 0.007504  | (0.0071, 0.0079) | <0.0001 |
| Weight²  | -0.000005 | (-0.00006, -0.00004) | <0.0001 |
| Height   | 0.2046    | (0.1097, 0.2995) | <0.0001 |
| Height*Weight | -0.00296 | (-0.00525, -0.00066) | 0.0116 |
| Height*Weight² | 0.000092 | (0.000058, 0.00013) | <0.0001 |
Table S3. Model removing potential outliers.

| Variable      | Estimate | 95% CI       | P-value |
|---------------|----------|--------------|---------|
| Intercept     | 5.7395   | (5.7112, 5.7677) | <0.0001 |
| Age           | 0.006205 | (0.0052, 0.0072)  | <0.0001 |
| Sex           |          |              |         |
| Male          |         |              |         |
| Female        | Ref     | (-0.1380, -0.0975) | <0.0001 |
| Weight        | 0.007502 | (0.0071, 0.0079)  | <0.0001 |
| Weight^2      | -0.00005 | (-0.00006, -0.00004) | <0.0001 |
| Height        | 0.2160   | (0.1265, 0.3056)  | <0.0001 |
| Height*Weight | -0.00327 | (-0.00544, -0.0011) | 0.0031  |
| Height*Weight^2| 0.000088 | (0.000057, 0.00012) | <0.0001 |
We estimated a multivariable linear regression model to predict mean heart weight and the upper 95% confidence interval limit using data on autopsies performed at the Cook County Medical Examiner’s Office (2014-2017). A cross-validation method was used to build the model, where half of the data were randomly selected and used to build the model and the other half were used to test the model. Common model selection criteria including Akaike information criterion (AIC), Bayesian information criterion (BIC), and cross-validated mean squared error (CV-MSE) were used to evaluate and select the final set and form of independent variables. Variance inflation factors (VIF) were also considered to identify issues with multicollinearity.

Specifically, we began with a series of models considering age, sex, weight, height, body mass index (weight (kg)/height (m²)), and the interaction between weight and height (Table S1, Step 1). Continuous variables were centered at their respective means. Residuals were assessed after fitting the initial model including age, sex, weight, and height as independent variables. A log transformation was applied due to concerns with normality of residuals. We have included histograms of the model residuals from the original scale heart weight model (Figure S1, left) and log transformed heart weight model (Figure S1, right). Subsequent models were fit using the log transformed outcome.

Figure S1. Histograms of Model Residuals from Step 1.

Models with highly correlated independent variables, as identified by a variance inflation factor greater than 10, were removed from consideration. The remaining model with the lowest AIC, BIC, and CV-MSE was selected. As descriptive scatterplots (Figure S2) indicated, potentially higher-order relationships between age, weight, height and the heart weight outcome existed. Therefore, we then considered adding higher-order terms.
We considered age$^2$, height$^2$, and weight$^2$ in an iterative process (Table S1, Step 2). Further investigation of interactions with age (Table S1, Step 3) were performed to assess whether the effects of these factors on heart weight may be different for different ages. Consideration of interactions with age, which may warrant stratification by age group, was further explored with visualizations of associations by age group (<20 vs 20+). Scatterplots overlaid with Loess curves and boxplots did not provide suggest different effects of height, weight, or sex or heart weight for different age groups (Figures S3 and S4). Furthermore, models incorporating interactions with age did not improve model fit, but caused additional concerns with multicollinearity. Thus, stratification of models by age group and interactions with age were not reported.

**Figure S3. Scatterplots of height and weight versus heart weight.**

Loess curves added, stratified by age < 20 years (red) and age ≥ 20 years of age or older (blue)

**Figure S4. Box plots of observed heart weight, stratified by sex and age.**
The final model was ultimately selected by the lowest AIC, BIC, and CV-MSE, without violation of multicollinearity concerns (Table S1, Model 7). Residuals were also assessed for the final model (Figure S5).

**Figure S5. Histogram of Residuals from Final Model.**

From the final selected model, we established an equation for estimating the conditional mean heart weight and corresponding upper 95% confidence interval limit based on the estimated coefficients and standard error (Table S2).

Equation for upper 95th limit of heart weight (Equation 1 from manuscript)

\[ e^{2.88 - 0.12 \times female + 0.0065 \times age + 1.09 \times h + 0.047 \times w - 0.018 \times h \times w + 0.000092 \times h^2 - 0.0002 \times w^2 + 0.25} \]

Where \( age \) = age at death in years; \( h \) = (body height in meters – 1.72), which is the measured height minus the mean sample height; \( w \) = (body weight in kilograms – 83.48), which is the measured weight minus the mean sample weight; and \( female \) = 1 if female and 0 if male. The value 0.25 is used because it is the standard error of the estimate multiplied by 1.65.

Finally, a sensitivity analysis considered exclusion of potential outliers, as identified by large residuals, < -3 or > 3 (n=15); however model results remained consistent (Table S3). We then applied the final equation to the NSDC cohort. Observed heart weights above the upper 95% confidence interval limit, conditional on age, sex, weight, and height, were defined as cardiomegaly.