

Mathematical details of the stochastic HSC model

We describe hematopoiesis as a two-compartment model. The first compartment contains the HSC, while the second contains the clones (i.e., the entire production) of committed stem cells. In the first compartment we allow cells to make stochastic decisions to divide, specialize or die, with fates depending on the type of event. Thus, an HSC can experience a birth (division with rate λ ; when the maximum number of niches is reached, one daughter cell dies immediately), death (apoptosis at rate α) and emigration (through commitment with rate ν). The second compartment gets input from committed stem cells, and the duration of productive life of a clone is modeled exponentially (clonal death with rate μ).

In each compartment we consider separately offspring expressing the father's and the mother's phenotype. Thus, the total number of cells/clones in each compartment is a sum of two processes that differ only in phenotype. There is a restriction in the first compartment, in that when the total number of HSC exceeds the upper limit K , only one daughter cell survives.

More formally, we write $X_m(t)$ for the number of HSC expressing the mother's phenotype at time t , and $X_f(t)$ for the number of HSC expressing the father's phenotype. Let $Y_m(t)$ and $Y_f(t)$ be the corresponding numbers of clones. Then we have the following possible state changes, writing $X_\bullet(t) = X_m(t) + X_f(t)$, $Y_\bullet(t) = Y_m(t) + Y_f(t)$ and letting dt be an infinitesimally small time interval:

$$X_m(t + dt) = X_m(t) + 1 \text{ with probability } X_m(t)\lambda dt \text{ provided } X_\bullet(t) < K$$

$$X_f(t + dt) = X_f(t) + 1 \text{ with probability } X_f(t)\lambda dt \text{ provided } X_\bullet(t) < K$$

$$X_m(t + dt) = X_m(t) - 1 \text{ with probability } X_m(t)\alpha dt$$

$$X_f(t + dt) = X_f(t) - 1 \text{ with probability } X_f(t)\alpha dt$$

$$X_m(t + dt) = X_m(t) - 1 \text{ and } Y_m(t + dt) = Y_m(t) + 1 \text{ with probability } X_m(t)\nu dt$$

$$X_f(t + dt) = X_f(t) - 1 \text{ and } Y_f(t + dt) = Y_f(t) + 1 \text{ with probability } X_f(t)\nu dt$$

$$Y_m(t + dt) = Y_m(t) - 1 \text{ with probability } Y_m(t)\mu dt$$

$$Y_f(t + dt) = Y_f(t) - 1 \text{ with probability } Y_f(t)\mu dt$$

Thus, the first compartment is a bounded two-dimensional linear birth-death-emigration process, while the second compartment is a two-dimensional immigration-death process, where the immigration consists of the emigration from the first compartment, and the death is a linear process. See Guttorp (1995, Chapter 3; reference 5 of text) for the terminology used here.

Alternative depiction of the Montreal data set

Figure S1A

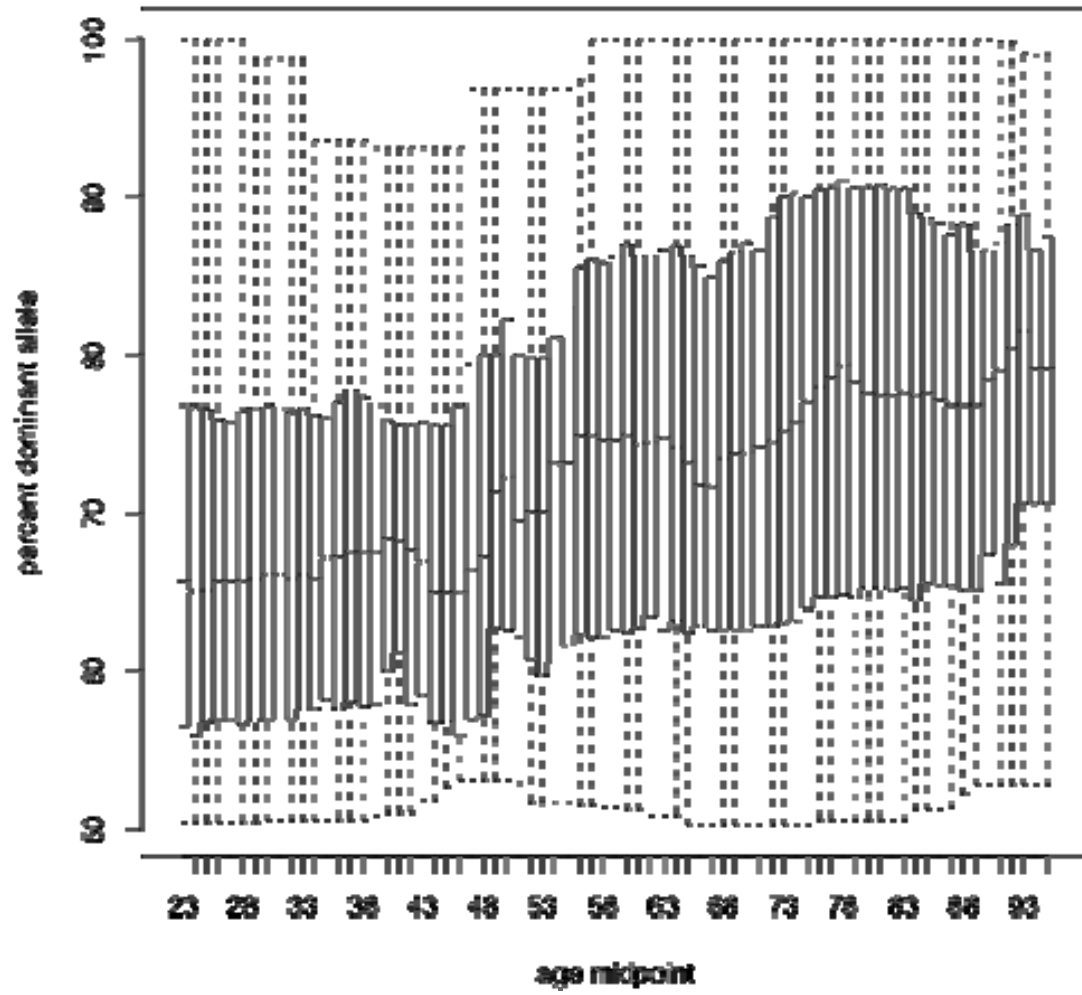


Figure S1B

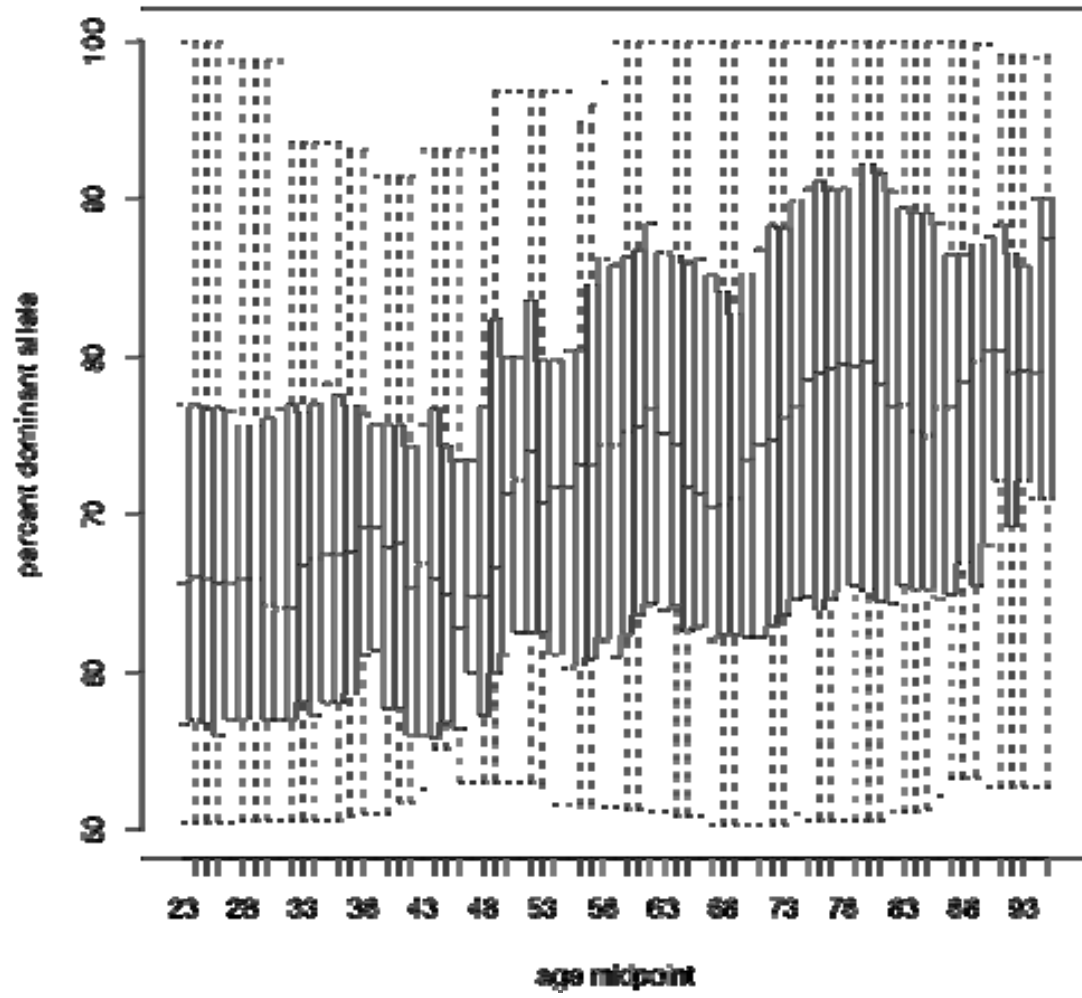


Figure S1C

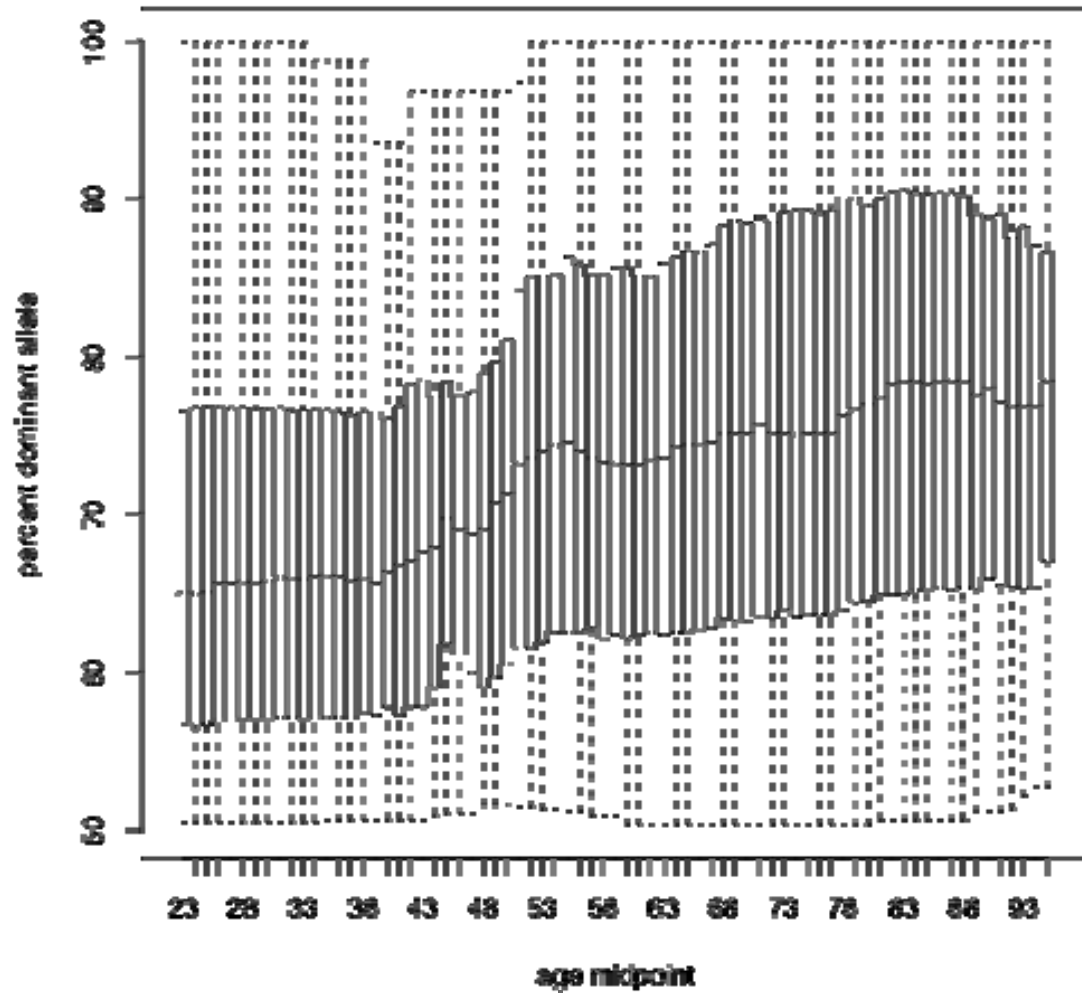


Figure S1

(A) depicts sequential boxplots of the data used in our calculations of the means and percentages of individuals with skewing (bin widths ± 5 complied at each age). (B) and (C) show the data that would have resulted at each age if a bin width of ± 3 and ± 10 , respectively, were instead used. A bin width of ± 5 was chosen as a trade-off between accuracy of age characterization and variability of the calculated statistics.