Mathematical supplement to "A Population-Based Job Exposure Matrix for Power-Frequency Magnetic Fields" by J.D. Bowman, J.A. Touchstone, and M.G. Yost. *J. Occup. Environ. Hyg.* Volume 4, Issue 9, 2007.

Proofs of Formulas for Summary Statistics

I. Arithmetic Summary Statistics (Eqs. 3–4 in main paper)

The arithmetic mean and standard deviation for the i-th occupational category (SOC or BOC) from published summary statistics from the data sources j are defined as:

$$AM_{ij} = \frac{\sum_{k=1}^{N_{ij}} x_{ijk}}{N_{ij}} \quad \text{(S1)}$$

$$SD_{ij}^2 = \frac{\sum_{k=1}^{N_{ij}} \left(x_{ijk} - AM_{ij}\right)^2}{N_{ij} - 1} = \frac{\sum_k x_{ijk}^2 - N_{ij} AM_{ij}^2}{N_{ij} - 1} \quad \text{(S2)}$$

here x_{ijk} is the TWA magnetic field for subject k and N_{ij} is the number of subjects from that source. (For spot measurements, x_{ijk} is the arithmetic mean of the spot measurements.)

As stated in the main paper (Eqs. 3–4), the summary arithmetic statistics for occupation i can be calculated exactly:

$$AM_i = \frac{\sum_j N_{ij} AM_{ij}}{N_i} \quad \text{(S3)}$$

$$SD_i^2 = \frac{\sum_j \left[\left(N_{ij} - 1\right) SD_{ij}^2 + N_{ij} AM_{ij}^2\right] - N_i AM_i^2}{N_i - 1} \quad \text{(S4)}$$

where $N_i = \sum_j N_{ij}$.

<u>Proof:</u> In terms of the original data, the mean is defined by:

$$AM_i \equiv \frac{\sum_{jk} x_{ijk}}{N_i} \quad \text{(S5)}$$

Then, substituting eq. S1 into Eq. S5 gives the desired result (Eq. S3).

Likewise, the occupational standard deviation is defined by:

$$SD_i^2 \equiv \frac{\sum_{jk}\left(x_{ijk} - AM_i\right)^2}{N_i - 1} = \frac{\sum_{jk} x_{ijk}^2 - 2AM_i \sum_{jk} x_{ijk} + N_i AM_i^2}{N_i - 1} \qquad \text{(S6)}$$

To obtain the sums in the numerator above, rearrange Eqs. S2 and S5:

$$\sum_k x_{ijk}^2 = \left(N_{ij} - 1\right) SD_{ij}^2 + N_{ij} AM_{ij}^2$$

$$\sum_{jk} x_{ijk} = N_i AM_i$$

and substitute these results into Eq. S6:

$$SD_i^2 = \frac{\sum_j \left[\left(N_{ij} - 1\right) SD_{ij}^2 + N_{ij} AM_{ij}^2\right] - 2N_i AM_i^2 + N_i AM_i^2}{N_i - 1}$$

$$= \frac{\sum_j \left[\left(N_{ij} - 1\right) SD_{ij}^2 + N_{ij} AM_{ij}^2\right] - N_i AM_i^2}{N_i - 1}$$

Q.E.D.

II. Geometric Summary Statistics (Eqs. 5–6)

In order to obtain GM_i and GSD_i, apply the results for arithmetic statistics to log-transformed data: $y_{ijk} = \ln x_{ijk}$ and the published geometric statistics:

$$\ln GM_{ij} = AM_{\ell ij} = \frac{\sum_k y_{ijk}}{N_{ij}} \qquad \text{(S7)}$$

$$\ln^2 GSD_{ij} = SD_{\ell ij}^2 = \frac{\sum_k \left(y_{ijk} - AM_{\ell ij}\right)^2}{N_{ij} - 1} \qquad \text{(S8)}$$

Applying Eq. S3 to the log-transformed data, we get the occupational geometric mean:

$$AM_{\ell i} \equiv \ln GM_i = \frac{\sum_j N_{ij} AM_{\ell ij}}{N_i}$$

The substitution of Eq. S7 gives Eq. 5 in the main paper:

$$\ln GM_i = \frac{\sum_j N_{ij} \ \ln GM_{ij}}{N_i} \tag{S9}$$

Likewise, applying Eq. S4 to Eq. S8 gives the occupational geometric standard deviation:

$$SD_{\ell i}^2 \equiv \ln^2 GSD_i = \frac{\sum_j \left[\left(N_{ij} - 1 \right) SD_{\ell ij}^2 + N_{ij} AM_{\ell ij}^2 \right] - N_i AM_{\ell i}^2}{N_i - 1}$$

or

$$\ln^2 GSD_i = \frac{\sum_j \left[\left(N_{ij} - 1 \right) \ln^2 GSD_{ij} \ + N_{ij} \ln^2 GM_{ij} \right] - N_i \ln^2 GM_i}{N_i - 1} \tag{S10}$$

which is Eq. 6 in the main paper.

In conclusion, Eqs. S3, S4, S9, and S10 are the formulas used to combine the published statistics in the JEM and to calculate statistics for higher level SOC codes. The same formulas are also used to combine published statistics with single measurements for each worker, for which N_{ij} =1, AM_{ij} = GM_{ij} = TWA or mean of spot measurements, SD_{ij} =0 and GSD_{ij} =1.

III. Weighted Ssummary Statistics (Eqs. 7–10)

For estimates that combine statistics from several occupations, we use weighted averages (Eq. 7 in the main paper):

$$AM_w = \frac{\sum_i w_i AM_i}{\sum_i w_i} \tag{S11}$$

where:

- w_i=1 when combining multiple “close categories” to estimate the exposure for a category without measurements.

- w_i = the number of U.S. employees in the SOC job when combining multiple SOC categories i into a single BOC.

The other weighted average statistics (Eqs. 8–10 in main paper) are then:

$$SD_w^2 = \frac{\sum_i w_i \left[\frac{N_i - 1}{N_i} SD_i^2 + AM_i^2 \right]}{\sum_i w_i} - AM_w^2 \quad \text{(S12)}$$

$$\ln GM_w = \frac{\sum_i w_i \ \ln GM_i}{\sum_i w_i} \quad \text{(S13)}$$

$$\ln^2 GSD_w = \frac{\sum_i w_i \left[\frac{N_i - 1}{N_i} \ \ln^2 GSD_i \ + \ \ln^2 GM_i \right]}{\sum_i w_i} - \ \ln^2 GM_w \quad \text{(S14)}$$

<u>Proof:</u> By substituting Eq. S5 into Eq. S11 and using:

$$\sum_{jk} w_i / N_i = w_i \, , \quad \text{(S15)}$$

AM_w can be written in terms of the individual data points:

$$AM_w = \frac{\sum_{ijk} \frac{w_i}{N_i} x_{ijk}}{\sum_{ijk} w_i / N_i} \quad \text{(S16)}$$

Following this example, we define the weighted standard deviation as:

$$\begin{aligned} SD_w^2 &\equiv \frac{\sum_{ijk} \frac{w_i}{N_i} \left(x_{ijk} - AM_w \right)^2}{\sum_{ijk} w_i / N_i} \\ &= \frac{\sum_{ijk} \frac{w_i}{N_i} x_{ijk}^2 - 2 AM_w \sum_{ijk} \frac{w_i}{N_i} x_{ijk} + AM_w^2 \sum_{ijk} \frac{w_i}{N_i}}{\sum_{ijk} w_i / N_i} \end{aligned} \quad \text{(S17)}$$

Then, substitute rearranged versions of Eqs. S16 and S6:

$$\sum_{ijk} \frac{w_i}{N_i} x_{ijk} = AM_w \sum_{ijk} w_i / N_i$$

$$\sum_{jk} x_{ijk}^2 = \left(N_i - 1\right) SD_i^2 + N_i AM_i^2$$

into Eq. S17 to get:

$$SD_w^2 = \frac{\sum_i \frac{w_i}{N_i} \left[\left(N_i - 1\right) SD_i^2 + N_i AM_i^2 \right]}{\sum_{ijk} w_i / N_i} - AM_w^2$$

Using Eq. S13, this simplifies to the desired result (Eq. S12). Q.E.D.

The weighted geometric statistics (Eqs. S13 and S14) are derived by the same method used with the occupational statistics GM_i (Eq. S9) and GSD_i (Eq. S10).