APPENDIX E-1

Spherical approximation: Equivalent circular diameters from two-dimensional profiles
Our method relies on treating microinfarcts as approximately spherical objects to enable estimation of the average microinfarct volume. The first step of this estimation is illustrated in Supplemental Figure 1: areas of 2-D microinfarct profiles are computed using image processing software after manually outlining their borders. These areas are then substituted into the formula for the area of a circle, which is then solved for the corresponding diameter of a circle ($d = 2 \cdot \sqrt{Area/\pi}$). This diameter is subsequently converted into an estimate of the equivalent volume of a sphere using a correction factor (see next section). This process is repeated for all microinfarcts, and the volume estimates are aggregated to calculate an average microinfarct volume. This method could fail to produce reasonable volume estimates under two general though implausible circumstances. First, if microinfarct shapes were systematically elongated either parallel or orthogonally to microtome sampling planes, this could result in exaggerated or reduced profile areas, leading in turn to over- or underestimation of average microinfarct volume, respectively. However, this problem does not arise under the reasonable assumption that the sampling plane orientation is random relative to microinfarct orientation within tissue. Second, certain unusual hypothetical microinfarct shapes (long thin radial extensions, for example) could compromise the validity of our method, but we are unaware of evidence suggesting such unusual shapes. Barring these possibilities, the proposed method is expected to produce reasonable estimates of the average microinfarct volume, a key parameter in our analysis.

Method for estimating the mean radius of spheres from random measurements
Consider a sphere with radius ρ. We obtain sample radius measurements by making vertical cuts at random heights, h, with all heights between 0 and 2ρ occurring with equal probability, i.e. $p(h) = 1/(2\rho)$. The average of these measurements, $\bar{r}$. will be smaller than the true radius, because most slices occur off-center.

To estimate the true radius, we need a correction factor. The relation between slice height h, measurement radius r, and true radius ρ is $h = \rho - \sqrt{\rho^2 - r^2}$ (Supplemental Figure 3). From this we can find a formula for the probability distribution over measurement values $p(r)$ using the standard probability transformation[1]

$$p(r) = p(h(r)) \left| \frac{\partial h(r)}{\partial r} \right|$$

which yields

$$p(r) = \frac{r}{\rho\sqrt{\rho^2 - r^2}}$$

Next, we can compute the expected value (mean) of the measurements:

$$\bar{r} = E[r] = \int_0^\rho r p(r) dr = \frac{\pi}{4}\rho = \int_0^\rho \frac{r^2}{\rho\sqrt{\rho^2 - r^2}} dr = \frac{\pi}{4}\rho.$$

Thus, given a set of measurements and the associated mean $\bar{r}$, we estimate of the true radius as

$$\hat{\rho} = \frac{4}{\pi}\bar{r}$$

If our sample contains spheres with a variety of different radii, we can estimate the true value based on the observed values via the law of iterated expectation[2]:

$$\bar{r} = E[r] = E[E[r|\rho]] = E\left[\frac{4\rho}{\pi}\right] = \frac{\pi}{4}\bar{\rho}.$$

Supplemental Figure 4 illustrates these concepts. For a range of true radius values (0 through 1000 microns) $\bar{\rho}$, a set of 25 sample observed radii values are plotted (drawn from a normal distribution with mean $\bar{\rho}$, and variance equal to 5% of the mean). The red "noisy" line shows the sample mean, whereas the blue line shows the true mean (ultimately obtained in the limit of a very large set of sample measurements).

Estimated mean radius for cerebral microinfarct data

We estimated the mean microinfarct radius using the method above, from a sample of 21 microinfarct specimens, measured by one of the authors (JAS). These values ranged from 43-363 microns, with a sample average value of $\bar{r} = 123$ microns. Using the correction factor above, the estimated mean microinfarct radius was $\hat{\rho} = \left(\frac{4}{\pi}\right)123 \approx 156.6$, hence the estimated mean diameter was ~313.2 microns. This estimates is further refined in the next subsection.

Correction when there is a lower limit of detection

The formula derived above requires revision if there is a lower limit to the size of detectable radii, conservatively estimated as~40 microns. The correction subtracts the portion of the integral for values of r between 0 and 40 microns

$$\delta(\rho) = \int_0^{40} rp(r)dr = \frac{\rho}{2}\tan^{-1}\left(\frac{40}{\sqrt{\rho^2 - 40^2}}\right) - \frac{40}{2\rho}\sqrt{\rho^2 - 40}$$

Given the observed sample mean $\bar{r}$, the corrected estimate for the true underlying mean radius can be obtained by solving the following expression for ρ,

$$\bar{r} = \frac{\pi}{4}\rho - \delta(\rho)$$

Setting $\bar{r} = 123$ microns, we obtain for our estimate $\hat{\rho} = 157.72$ microns, or diameter $d = 315.4$ microns. Thus, in our case, the correction factor is small, $\delta(\hat{\rho}) \approx 0.87$ microns.

Width of the 90% confidence interval vs number of pathology slides examined

The majority of neuropathology examinations reported in this work involved 9 sections. The confidence intervals for estimates of total brain microinfarct burden based on small tissue samples such as these are necessarily wide. Supplemental Figure 5 illustrates how the width of the confidence interval for estimates of total brain microinfarct number depends on the number of samples examined. This figure assumes that maximum likelihood estimate (MLE) of total microinfarcts burden is close to 1000 microinfarcts, while varying the number of pathology specimens N examined in arriving at this estimate. (Note that the solid curve is not a strictly decreasing function of N, because the MLE cannot always be precisely equal to 1000 in our model). The confidence interval width decreases from 2454 for estimates based on only 9

sections, to 476 after 100 sections are examined. The fitted curve (dashed line) shows that the confidence interval width decreases at a rate that is approximately inversely proportional to the square root of the number of pathology sections examined.

References

1. Bishop CM. Pattern recognition and machine learning. New York: Springer; 2006.
2. Papoulis A. Probability, random variables, and stochastic processes. 2nd ed. New York: McGraw-Hill; 1984.

Fig e-1

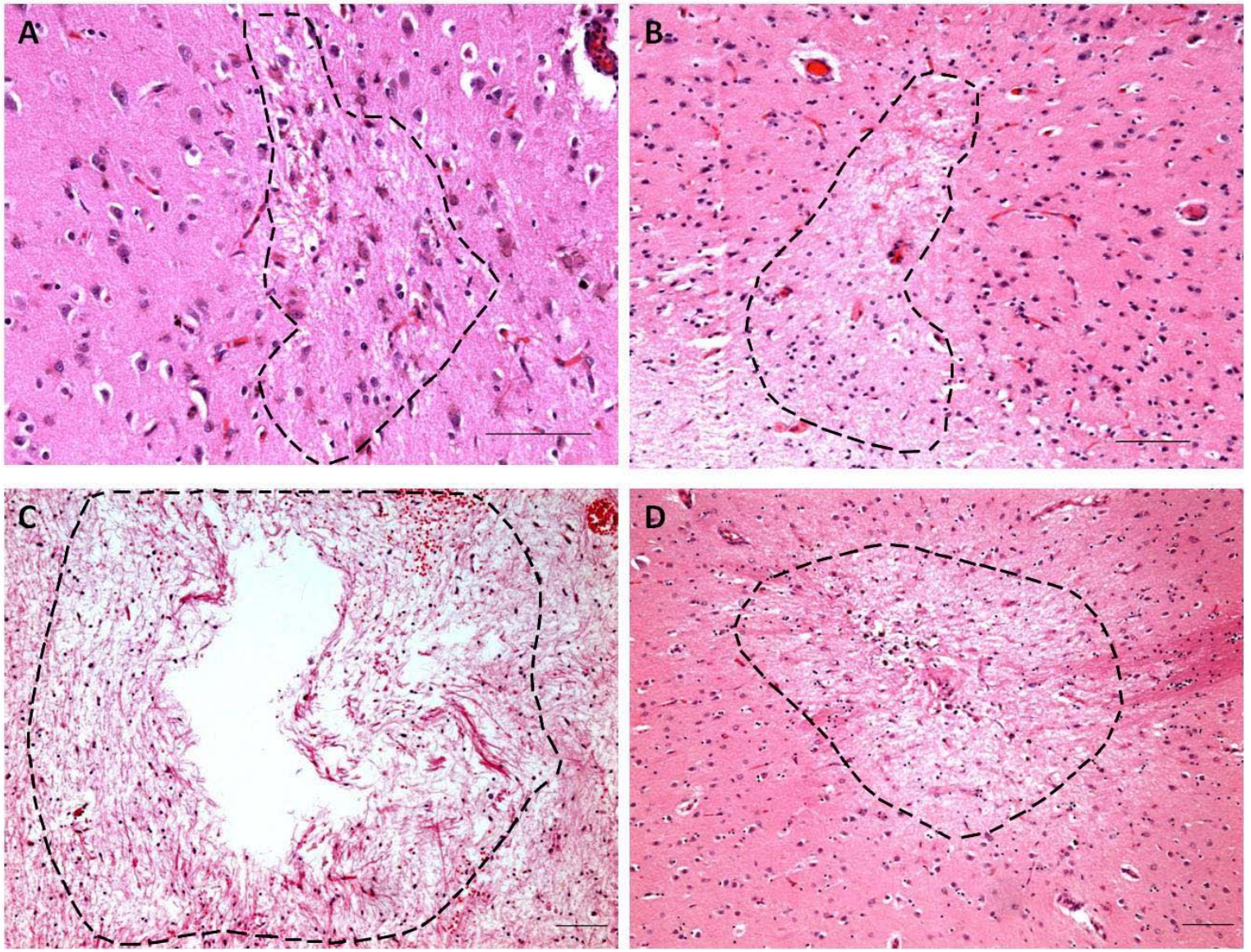

Fig e-1: Photographed examples of typical microinfarct profiles. A. Microinfarct with slight central cavitation in the middle temporal cortex (diameter: 303.96 µm) B. Microinfarct with tissue pallor without cavitation in the superficial midfrontal cortex (diameter: 298.65 µm). C. Cystic microinfarct in midfrontal cortex (diameter: 856.78 µm). D. Microinfarct with old focal central slit hemorrhage in the head of the caudate (diameter: 643.91 µm). Dashed lines encompass the region of microinfarction defined by cell loss with cavitation and/or tissue pallor and bordered by surrounding gliosis. Hematoxylin and Eosin stain. Bar width = 100 microns.

Fig e-2

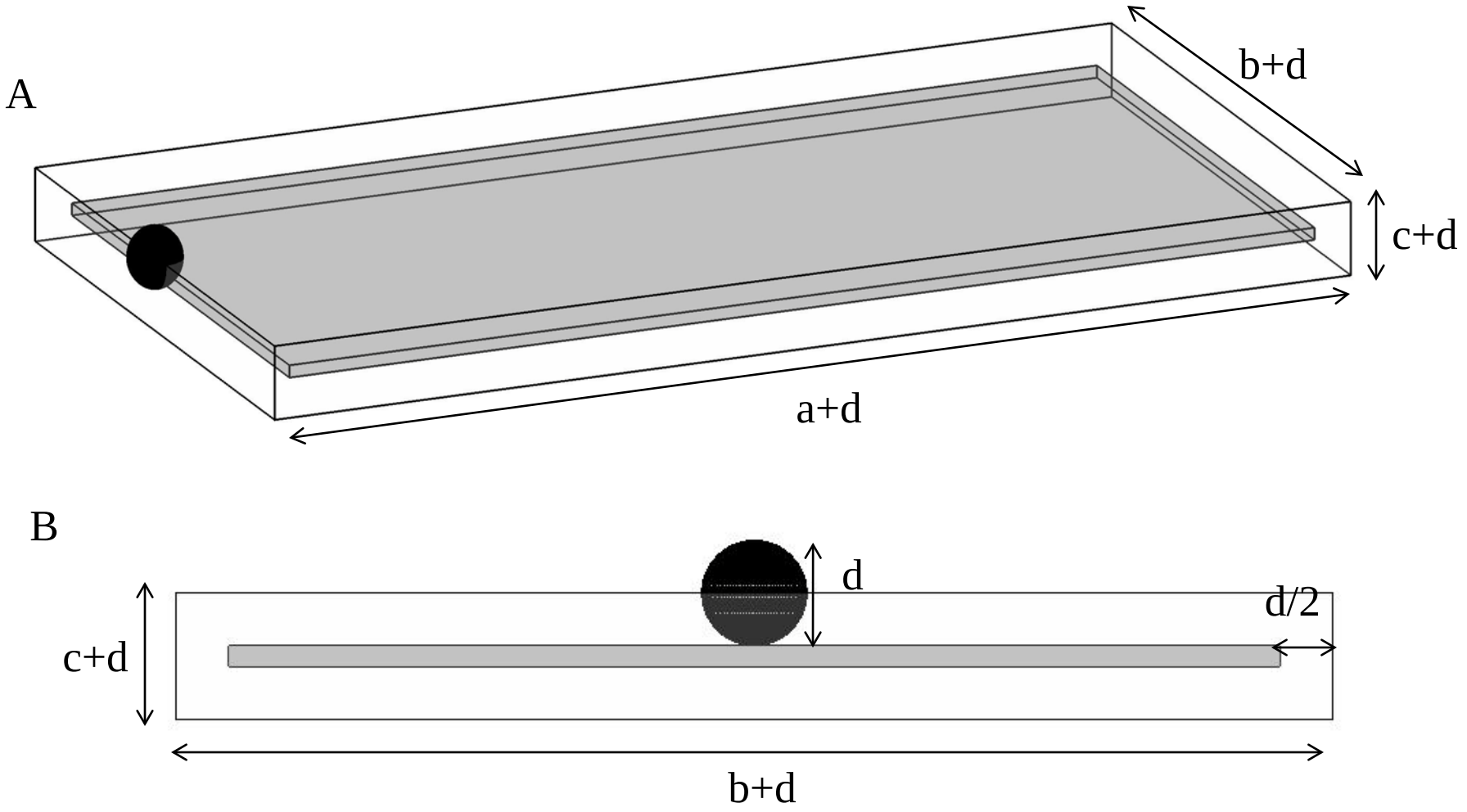

Fig e-2. **Schematic of tissue slice and micro-infarct geometry.** A. The gray slab represents a microtome section of brain having width , length , and thickness . The filled sphere is a micro-infarct with diameter, d. The larger open slab represents the "catchment" area for potential intersection of at least a portion of a micro-infarct with diameter d. Thus, its dimensions are width a+d, length b+d, and thickness c+d. B. Side view of panel A, to illustrate the catchment area bounded by half the diameter of micro-infarct "extra" around the edges of the gray slab of actual tissue under review.

Fig e-3

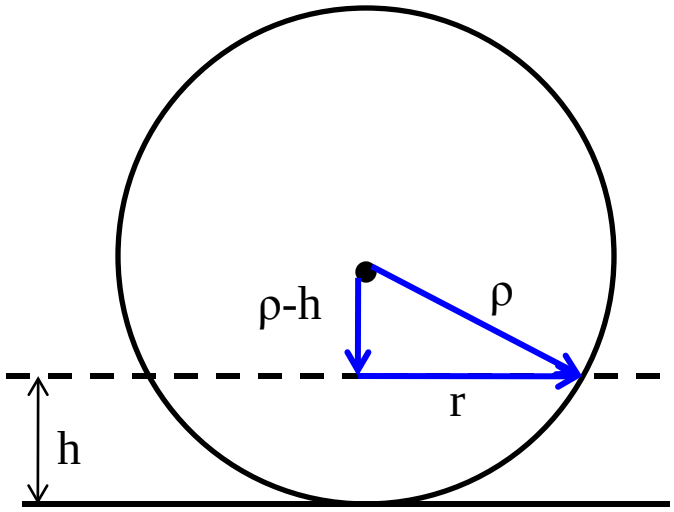

Fig e-3. Schematic of the relationship between the true and observed radius of a micro-infarcts seen in a microtome slice.

Fig e-4

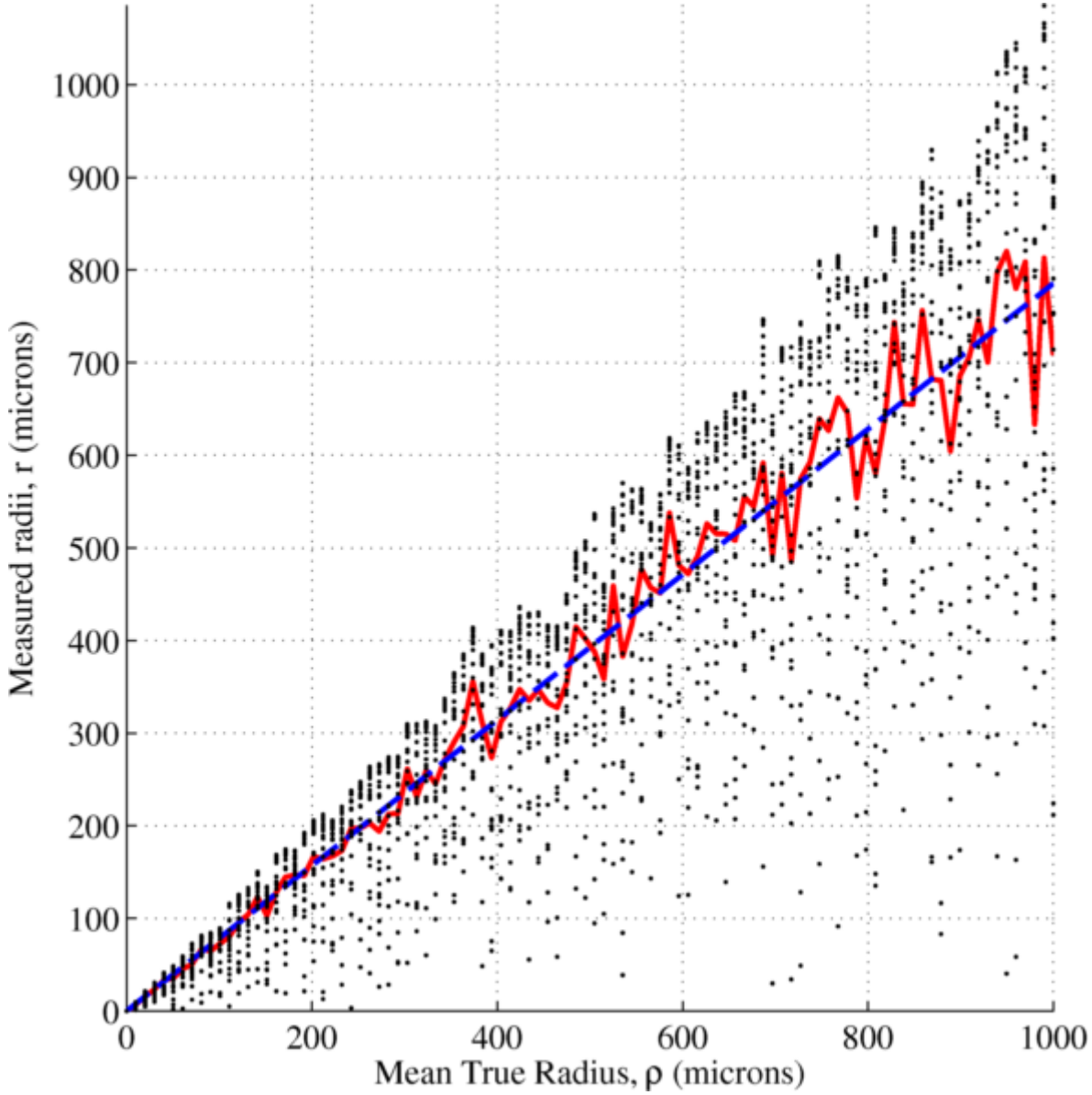

Fig e-4. Scatter plot showing the relation of observed versus actual micro-infarct radii. For each true radius value (0-1000 μM), we plot a distribution of 25 possible observed radii (dots). The red line shows the mean value of these 25 values at each X-axis point; the blue line is the expected true mean. Note all values fall below the identity line (which is not shown), as expected given that the observed values are generally smaller than the true value due to the detection limits discussed in the main text.

Fig e-5

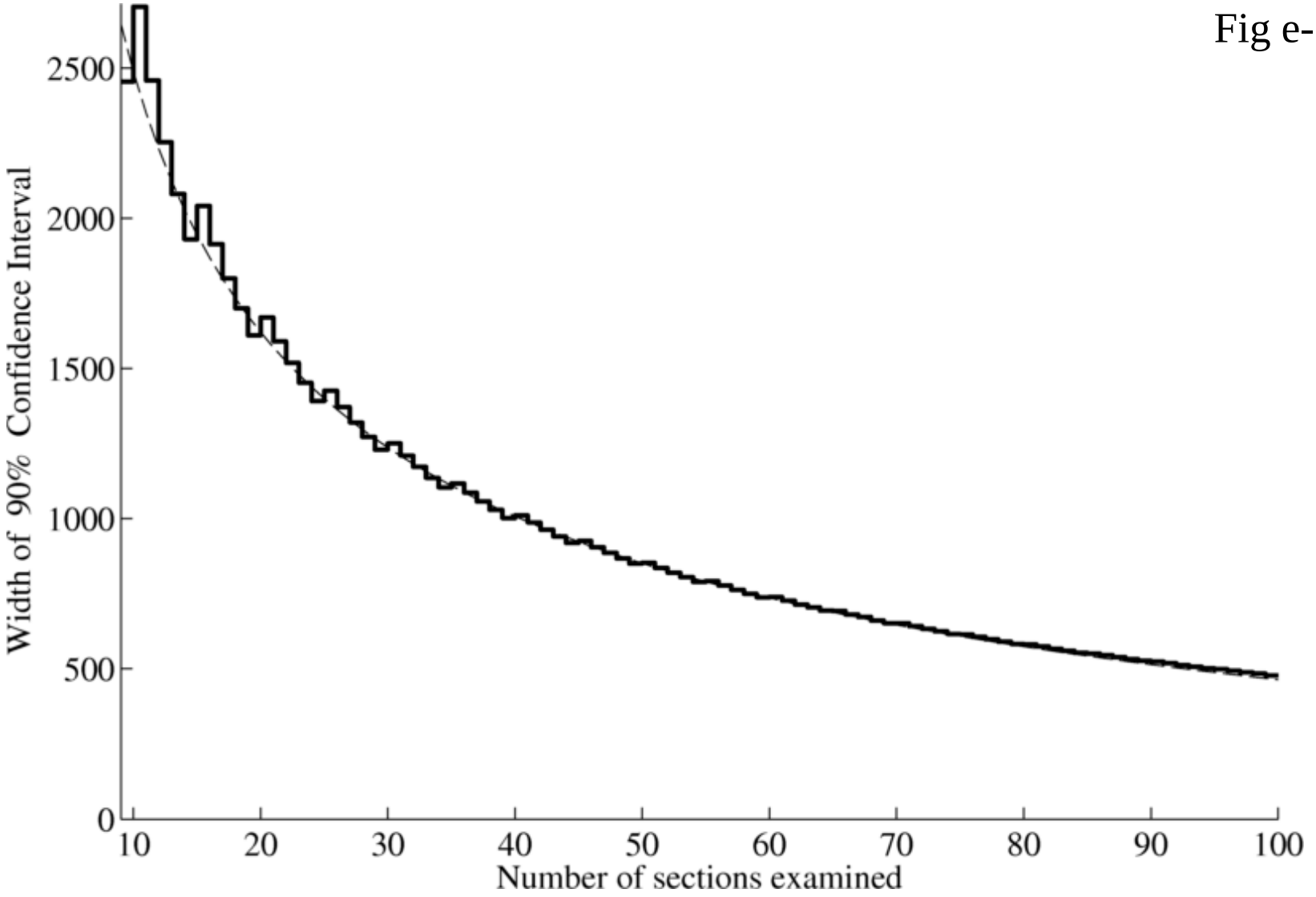

Fig e-5. Width of the 90% confidence interval for the estimated total brain microinfarct volume vs. number of pathology slides examined. Assuming a maximum likelihood estimate (MLE) of of microinfarcts and taking as the independent variable (x-axis) the number of samples upon which the MLE is based, we plot on the y-axis the width of the 90% confidence interval (solid line). The decrease in a fitted curve of the form is shown as a dashed line.