Appendix 1. Diagnostic criteria for HIV-Associated Neurocognitive Disorders.

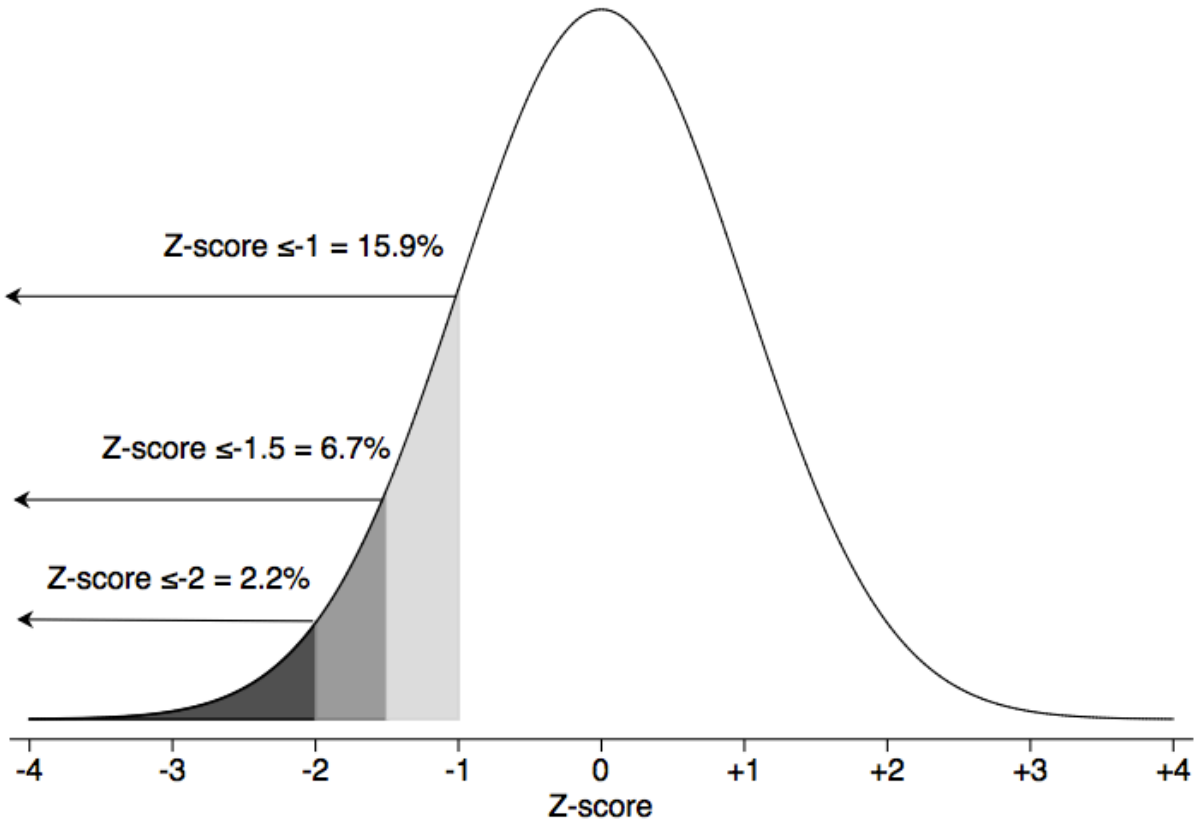
HIV-associated Asymptomatic Neurocognitive Impairment[1]	A. Acquired impairment in cognitive functioning with scores at least 1 standard deviation below the mean in at least two cognitive domains B. Does not interfere with everyday functioning
HIV-associated Mild Neurocognitive Disorder[1]	A. Acquired impairment in cognitive functioning with scores at least 1 standard deviation below the mean in at least two cognitive domains B. Mild interference in daily functioning
HIV-associated Dementia[1]	A. Marked acquired impairment in cognitive functioning with scores at least 2 standard deviations below the mean in at least 2 cognitive domains B. Marked interference in day-to-day functioning

* Other criteria include: (1) Does not meet criteria for delirium or dementia; (2) No evidence of pre-existing causes.

Reference

1. Antinori A, Arendt G, Becker J, Brew B, Byrd D, Cherner M, et al. Updated research nosology for HIV-associated neurocognitive disorders. Neurology. 2007;69:1789-99.

Appendix 2. The proportion of a normal population theoretically expected to have scores below certain threshold Z-scores on a single test assuming a normal distribution.



Appendix 3. Summary of formulas used for the calculation of bias and effective sample size.

Summarized from Bross I. Misclassification in 2 x 2 tables. International Biometric Society 1954;10:478-486.

Bias in Prevalence Estimate

Let:
n= number of individuals drawn from population
u= actual number of "problem" cases in sample (true positives=people with HAND, classified as having HAND)
v= number of "problem" individuals incorrectly classified as "non-problem" cases (false negatives=people with HAND, classified as no HAND)
w= number of "non-problem" individuals who are incorrectly classified as "problem" cases (false positives)
x= apparent number of problem individuals

Therefore:
$\mathbf{x=u-v+w}$

Let:
p = proportion of "problem" cases in the population; true prevalence
q = is the proportion of "non-problem" cases
θ = probability of misclassifying a "problem" individual, e.g. false negatives
ϕ = probability of misclassifying a "non-problem" individual, e.g. false positives

If misclassification exists, the estimated proportion of "problem" cases will not be p as above but instead:
$E(x/n) = p - p\theta + q\phi$ where $q = (1-p)$

Then:
$E(x) = n\,p\,K$ where $K=1- \theta + q/p\,(\phi)$

Thus:
Estimated prevalence = $E(x/n) = p\,K$

And:
Bias in prevalence estimate = $p\,K - p$

Effective sample size

The quantity K' measures the effect of misclassification on the variance of the estimate p where:

$K' = 1-2\phi - 2\theta + \theta/q + \phi/p$

The accepted measure of efficiency is:

Efficiency = $1/(K'+D)$ where $D= [(q\phi - p\theta)^2\, n]/pq$

And an accepted measure of effective sample size is:

Effective sample size= n * efficiency

Appendix 4. Tables underlying Figures 2 and 3.

Table 1. The proportion of a young Kenyan HIV-uninfected population who meet criteria for cognitive impairment using various definitions of impairment and various criteria for an abnormal domain (n=84).

Criteria for an abnormal domain:	**Z-scores ≤-1 in ≥2 domains** **% [95%CI]**	**Z-scores ≤-1.5 in ≥2domains** **% [95%CI]**	**Z-scores ≤-2 in ≥2 domains** **% [95%CI]**
One test in domain abnormal	48% [37%, 59%]	20% [11%, 29%]	6% [1%, 11%]
Average score in domain abnormal	20% [11%, 29%]	6% [1%, 11%]	0%
At least two tests in domain abnormal	19% [10%, 28%]	4% [-0.5%, 8%]	0%
All tests in domain abnormal	8% [2%, 14%]	1% [-1%, 4%]	0%

Table 2. The percentage of a simulated normal population who meet criteria for cognitive impairment using various definitions of impairment and various criteria for an abnormal domain.

	Z-scores ≤-1 in ≥2 domains				**Z-scores ≤-1.5 in ≥2 domains**				**Z-scores ≤-2 in ≥2 domains**			
Criteria for an abnormal domain:	**Emp**	**Low**	**Mod**	**High**	**Emp**	**Low**	**Mod**	**High**	**Emp**	**Low**	**Mod**	**High**
One test in domain abnormal	56	74	45	27	24	29	22	13	7	5	8	5
Average score in domain abnormal	30	31	25	20	11	7	11	9	3	1	4	3
At least two tests in domain abnormal	11	4	15	17	2	0	5	7	0	0	1	2
All tests in domain abnormal	8	2	11	15	1	0	3	6	0	0	1	2

Table 3. The percentage of a simulated normal population who meet criteria for cognitive impairment using neuropsychological test batteries of various sizes.

	Empiric Correlation (Within 0.31, Between 0.17)				**Moderate Correlation (Within 0.5, Between 0.5)**			
Number of Domains	**2 tests**	**3 tests**	**4 tests**	**5 tests**	**2 tests**	**3 tests**	**4 tests**	**5 tests**
Z≤-1								
2	4	5	5	5	8	9	10	10
3	10	10	11	11	14	14	14	15
4	15	16	16	16	17	18	18	18
5	21	21	21	21	21	20	20	20
6	26	25	25	25	23	23	22	22
7	30	30	29	29	26	25	24	23
8	34	34	33	32	28	26	25	24
9	38	37	36	36	30	27	26	25
10	41	40	39	38	31	29	27	26
Z≤-1.5								
2	1	1	1	1	3	3	3	4
3	3	3	3	3	5	5	6	6
4	4	5	5	5	7	7	7	7
5	6	7	7	7	9	8	8	8
6	8	9	9	9	10	10	9	9
7	10	11	11	11	11	11	10	10
8	12	13	13	13	12	12	11	11
9	14	14	15	15	13	12	12	12
10	16	16	16	16	14	13	13	12
Z≤-2								
2	0	0	0	0	1	1	1	1
3	0	1	1	1	1	2	2	2
4	1	1	1	1	2	2	2	2
5	1	2	2	2	3	3	3	3
6	2	2	2	2	3	3	3	3
7	2	3	3	3	4	4	3	3
8	3	3	3	3	4	4	4	4
9	3	4	4	4	5	4	4	4
10	4	4	4	5	5	5	4	4

Appendix 5. Tables supporting Figures 3 and 4.

Table 1. Bias in prevalence estimate as a function of the false positive rate and true prevalence of disease (selected values only).*

True Prevalence	False Positive Rate	Estimated Prevalence	Bias in prevalence estimate
0.02	0.01	0.03	0.01
	0.05	0.07	0.05
	0.1	0.12	0.10
	0.2	0.22	0.20
	0.4	0.41	0.39
0.1	0.01	0.10	0.00
	0.05	0.14	0.04
	0.1	0.19	0.09
	0.2	0.28	0.18
	0.4	0.46	0.36
0.25	0.01	0.25	-0.01
	0.05	0.28	0.03
	0.1	0.31	0.06
	0.2	0.39	0.14
	0.4	0.54	0.29
0.4	0.01	0.39	-0.01
	0.05	0.41	0.01
	0.1	0.44	0.04
	0.2	0.50	0.10
	0.4	0.62	0.22

* Assumes false negative rate of 0.05.

Table 2. Effective sample size as a function of the false positive rate, true prevalence of disease and sample size (selected values only).*

Actual Sample Size (n=100)			Actual Sample Size (n=1000)		
Prevalence	**False positive rate**	**Effective sample size**	**Prevalence**	**False positive rate**	**Effective sample size**
0.02	0.01	55	**0.02**	0.01	186
	0.05	7		0.05	8
	0.1	2		0.1	2
	0.2	0		0.2	1
	0.4	0		0.4	0
0.1	0.01	95	**0.1**	0.01	824
	0.05	32		0.05	52
	0.1	10		0.1	12
	0.2	3		0.2	3
	0.4	1		0.4	1
0.25	0.01	100	**0.25**	0.01	893
	0.05	72		0.05	227
	0.1	31		0.1	45
	0.2	9		0.2	10
	0.4	2		0.4	2
0.4	0.01	94	**0.4**	0.01	554
	0.05	95		0.05	702
	0.1	59		0.1	130
	0.2	19		0.2	23
	0.4	5		0.4	5