Medicalising normality?

Using a simulated dataset to assess the performance of different diagnostic criteria of cognitive impairment

J. Underwood[1], D. De Francesco[2], R.Leech[3], C. Sabin[2] and A.Winston[1], on behalf of the POPPY study group.

[1]Division of Infectious Diseases, Imperial College London, UK, [2]UCL - Royal Free Campus, Research Department of Infection & Population Health, London, UK, [3]Division of Brain Sciences, Imperial College London, UK

Background

The reported prevalence of cognitive impairment remains similar to that from the pre-antiretroviral therapy era ~50%.[1] This may be partially artefactual due to the diagnostic methods commonly used, as studies with HIV-negative control groups also reportedly have high rates of cognitive impairment (29-36%).[2,3]

Commonly used methods of identifying impairment rely on defining impairment in those exceeding a threshold of deviation from a normative score.

These normative scores are defined for each neuropsychological test by testing a range of healthy participants and accounting for demographics such as age and level of education, so that normative scores for all tests are normally distributed with a consistent mean and standard deviation (SD) – for T-scores 50 (10).

Aims

In this study, we evaluated the diagnostic performance of the HIV-associated neurocognitive disorder ('Frascati' criteria) and the global deficit score (GDS) methods in comparison to a novel, multivariate method of diagnosis which we outline here.

Methods

Normative dataset

- Firstly, a simulated '*normative*' dataset (n=10,000) was created with mean (SD) T-scores of 50 (10) across six simulated cognitive domains following a multivariate normal distribution.
- To make the model more accurate and to account for the covariance between cognitive domains, the inter-domain correlation coefficients were set to match those of the HIV-positive group from the Pharmacokinetic and Clinical Observations in PeoPle Over fiftY (POPPY) study (n=290), where participants underwent cognitive function testing using a computerised battery (CogState) covering six cognitive domains.

Multivariate assessment of cognitive impairment

- The Mahalanobis distance (analogous to a multivariate standard deviation) was calculated for each participant from the centre of the simulated *normative* dataset.
- Each Mahalanobis distance was given direction so that positive values represented scores in general above the mean (i.e. mean T-score [the 'global' score] > population mean [50]) and negative values represented scores below (i.e. impairment: global < 50).
- This value was then compared to a critical value to determine if impairment was present using the β distribution:[4]

$$critical\ value\ = -\sqrt{\frac{(n-1)^2}{n}\cdot\beta_{\alpha;\frac{p}{2},\frac{(n-p-1)}{2}}}$$

Where:
n = the number of subjects
p = the number of domains/tests
β = the critical value from the β distribution with parameters $\frac{p}{2}$ and $\frac{(n-p-1)}{2}$ with α=0.05 (i.e. corresponding to the bottom 5th percentile of a normative population)

Next, the apparent prevalence of cognitive impairment was determined by applying the Frascati and GDS methods of impairment as well as the novel, multivariate method (table 1).

Table 1. Summary of the different diagnostic criteria.

Diagnostic criterion	Definition of abnormal test result	Definition of abnormal cognitive domain	Affected domains
Frascati[5]	1 SD below normative mean	mean domain performance abnormal	≥ 2
Global deficit score (GDS)[6]	**GDS score ≥ 0.5** (the GDS score is obtained by converting domain T scores into deficit score ranging from 0 to 5 and then averaging them)		
Mahalanobis distance	Mahalanobis distance (analogous to a multivariate standard deviation taking account of performance in all cognitive domains simultaneous) exceeds a critical value		

Methods cont...

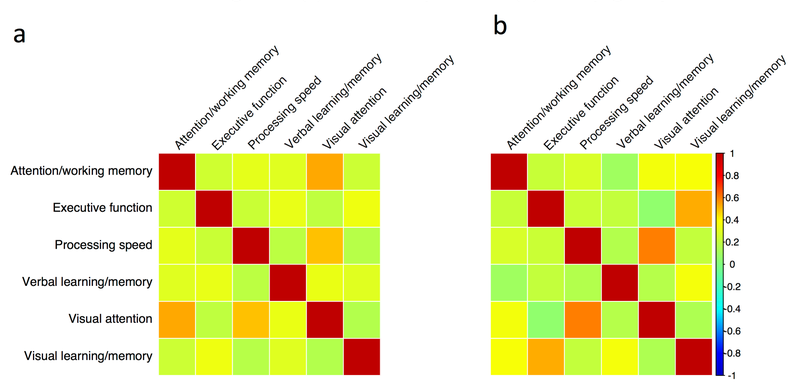
Test dataset

- Next, an '*impaired*' population was added to this hypothetical *normative* control population to create a '*test*' dataset with mean (SD) T-score of 30 (10).
- The prevalence of '*impairment*' was initially set at 10%.
- A study population (n=290) was then sampled from the *test* dataset and sensitivity, specificity, predictive values and accuracy were determined by comparing the subjects labelled as impaired by each method to the subjects defined *a priori* as *impaired* (true positives).
- Bootstrapping was then performed (10,000 replicates) to determine the mean prevalence of cognitive impairment and performance characteristics with 95% confidence intervals.
- As a final step true impairment was varied from 0-40% by increasing the size of the *impaired* sample and the previous steps were repeated.

Results

HIV-positive and HIV-negative groups from POPPY had comparable inter-domain correlation coefficients, justifying the use of the patient group's data to model the *normative* population (χ^2_{15}=19.5, p=0.19; figure 1).

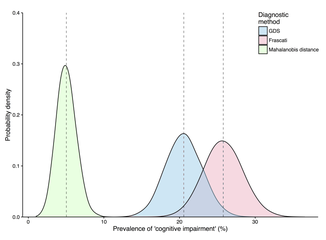
Figure 1. Comparison of the cognitive domain correlation matrices for the POPPY HIV-negative (a) and HIV-positive (b) study groups



The proportion of the *normative* population labelled as having cognitive impairment was significantly lower for the Mahalanobis distance method vs. Frascati or GDS (table 2).

Table 2 and Figure 2. The expected prevalence of cognitive impairment in a normative population

Diagnostic criterion	Mean prevalence (%)	95% CIs (%)
Frascati	25.8	21.7-30.0
GDS	20.6	16.9-24.5
Mahalanobis distance	5.0	3.1-7.2



Using the test dataset, where 10% had cognitive impairment, (figure 3a for a graphical illustration), the prevalence of cognitive impairment increased to 33.6% (29.3-38.3%) using the Frascati criteria (figure 3b); 28.5% (24.8-32.8%) for GDS (figure 3c); and 12.1% (9.0-15.1%) using the Mahalanobis distance method (figure 3d).

Figure 3. Histograms of the *test* dataset with a 10% prevalence of cognitive impairment.

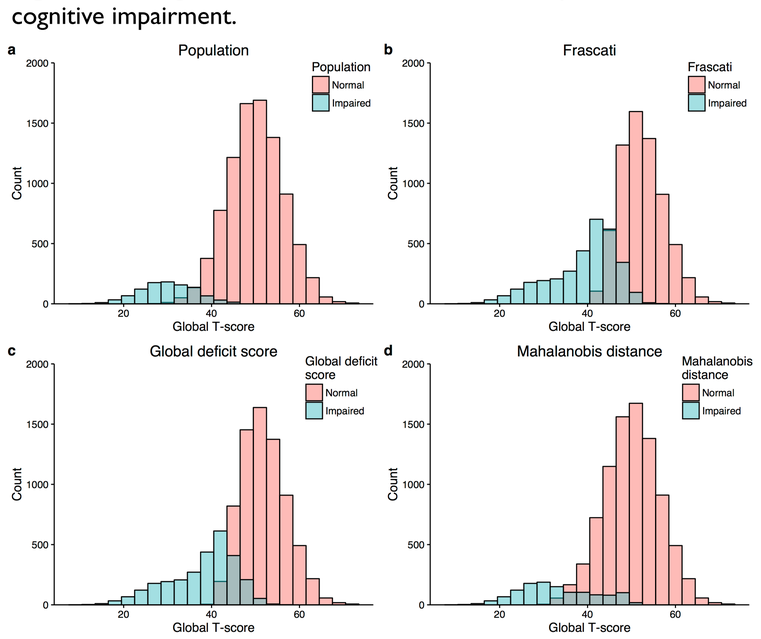


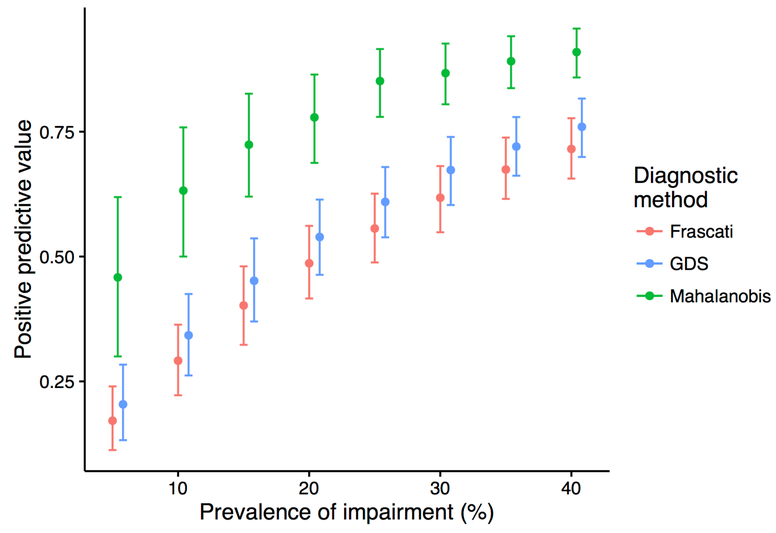
Table 3. Performance of three definitions of cognitive impairment in a simulated population with a 10% prevalence of cognitive impairment.

Diagnostic criterion	SENS (95% CI)	SPEC (95% CI)	PPV (95% CI)	NPV (95% CI)	Accuracy (95% CI)
Frascati	99% (95-100%)	74% (69-78%)	29% (22-37%)	100% (99-100%)	76% (72-80%)
GDS	98% (93-100%)	79% (74-83%)	34% (26-42%)	100% (99-100%)	81% (77-85%)
Mahalanobis distance	77% (65-89%)	95% (93-95%)	61% (48-74%)	97% (96-99%)	93% (90-95%)
Combination	77% (65-90%)	95% (93-97%)	65% (52-78%)	97% (96-99%)	94% (91-96%)

Results cont…

When the prevalence of impairment was 5% in the test dataset, the positive predictive value was below 25% for Frascati and GDS compared with nearly 50% for Mahalanobis distance method (figure 4).

Figure 4. Positive predictive value for each diagnostic method by prevalence of impairment.



Discussion

The Frascati and GDS methods classify over 20% of a *normative* control population as cognitively impaired. This may be partly responsible for the reportedly high prevalence of cognitive impairment observed in clinical studies.

Over diagnosis of cognitive impairment is potentially problematic as it may:

- Overstate the burden of disease.
- Lead to unnecessary investigations.
- Cause psychological distress.
- Make further investigation of the underlying pathophysiology more difficult.

The novel multivariate method outlined here was more accurate with a greater positive predictive value than Frascati and GDS.

Further research testing this method needs to be performed, specifically to determine if it identifies people with greater neuropathology.

For an interactive model of these data, please see: https://jonathan-underwood.shinyapps.io/cognitive_impairment_comparison/

Conclusion

The commonly used diagnostic criteria of HIV-associated cognitive impairment label a significant proportion of any population as cognitively impaired, with a substantial over-estimate of the true proportion. These findings have important implications for clinical research regarding cognitive health. More accurate methods of diagnosis should be implemented, with multivariate techniques offering a promising solution.

Lay summary

People living with HIV are more likely to be diagnosed with cognitive impairment than HIV-negative people. However, the amount of genuine cognitive impairment may be overstated because of the way 'impairment' is defined. Over diagnosis may be reduced with a new method – presented here.

References

1. Heaton RK *et al.* Neurology; 2010;75: 2087–2096
2. Su T *et al.* AIDS. 2015;29: 547–557..
3. McDonnell J *et al.* J Acquir Immune Defic Syndr. 2014;67: 120–127
4. De Maesschalck R *et al.* Chemometrics and Intelligent Laboratory Systems. 2000;50: 1–18 .
5. Antinori A *et al*, *Neurology* (2007)
6. Carey CL *et al. J Clin Exp Neuropsyc* (2004)

Acknowledgements

POPPY Management Team
Marta Boffito, Paddy Mallon, Frank Post, Caroline Sabin, Memory Sachikonye, Alan Winston

Scientific Steering Committee
Jane Anderson, David Asboe, Marta Boffito, Lucy Garvey, Paddy Mallon, Frank Post, Anton Pozniak, Caroline Sabin, Memory Sachikonye, Jaime Vera, Ian Williams, Alan Winston

POPPY Sites and Trials Unit
- Elton John Centre, Brighton and Sussex University Hospital (Martin Fisher, Amanda Clarke, Jaime Vera, Andrew Bexley, Celia Richardson)
- St Stephen's Centre, Chelsea and Westminster Hospital (Marta Boffito, David Asboe, Anton Pozniak, Chris Higgs, Elisha Seah, Stephen Fletcher, Michelle Anthonipillai, Ashley Moyes, Katie Deats)
- Homerton Sexual Health Services, Homerton University Hospital (Jane Anderson, Sifiso Mguni, Rebecca Clark, Rhiannon Nevin-Dolan)
- Caldecot Centre, King's College Hospital (Frank Post, Lucy Campbell, Selin Yurdakul, Sara Okumu, Louise Pollard)
- HIV Molecular Research Group, Mater Misericordiae University Hospital (Paddy Mallon, Alan Macken, Bijan Ghavani-Kia, Joanne Maher, Maria Byrne, Ailbhe Flaherty)
- Department of Infection and Population Health, University College London (Ian Williams, Damilola Otiko, Laura Phillips)
- St. Mary's Hospital London, Imperial College Healthcare NHS Trust (Alan Winston, Lucy Garvey, Matthew Scott, Linda McDonald)
- Imperial Clinical Trials Unit, Imperial College London (Andrew Whitehouse, Laura Burgess, Daphne Babalis)

Funders
The POPPY study is funded from investigator initiated grants from BMS, Gilead Sciences, Janssen, Merck and ViiV Healthcare.