Paper title:

“INSIDE THE ‘BLACK BOX’ OF THE ANTIBODY TEST: DECONSTRUCTING OFFICIAL CLASSIFICATION OF ‘RISK’ IN THE TEST ALGORITHMS USED FOR IDENTIFYING THE HUMAN IMMUNODEFICIENCY VIRUS.”

Author:

KEVIN CORBETT

Presented at:

Research Committee 23: Sociology of Science and Technology
Session 6: The Social and Ethical Implications of Biotechnology
Wednesday 26th July, 13.30-15.30pm
XVI World Congress of Sociology
Durban, South Africa
23-29 July 2006

Contact details:

Dr Kevin Corbett  RN BA (Hons) HDFA PGCE MSc PhD
Senior Research Fellow in Health and Social Care
Centre for Public Health
Faculty of Health and Applied Social Sciences
Liverpool John Moores University
70 Great Crosshall Street
Liverpool
England L2 3AB
United Kingdom

Tel: 0151 231 4474
Fax: 0151 231 4218
E-mail: K.Corbett@ljmu.ac.uk
Abstract:
This paper interrogates the last 20 years in the British experience of using official antibody test algorithms to detect the human immunodeficiency virus (HIV). Case definitions of the Acquired Immunodeficiency Syndrome (AIDS) cite antibody test methodologies licensed since 1985 for screening purposes and derived from laboratory identification of HIV. Two common (yet surrogate) methodologies are the enzyme-linked immunosorbent assay (ELISA) and the Western blot (WB), both used for screening human populations. Test manufacturers publicise the interpretative flexibility of these tests, which may produce false or indeterminate results, given laboratory identification of HIV is cited as problematic, time-intensive and as using surrogate techniques. Globally, public health officials publish differing algorithms for testing of human subjects. The paper shows how these algorithms (whilst aiming to balance test specificity/sensitivity), are based on perceptions of ‘risk’ of exposure determined during pre-test dialogue: how the test subject is positioned as ‘high’/‘low’ risk and within a hierarchy of exposure categories. The interpretation of indeterminate results is problematic given the possibility of false results, which are ruled out by estimating the risk of exposure (‘window period’) and the seroprevalence in the population of the test subject. It is argued that during the last 20 years experience with these test algorithms the interpretation of the test ‘result’ is not wholly ‘objective’ or laboratory-determined, as it relies as much upon the classification of the test subject as being ‘at risk’ during pre-test dialogue as it does upon the “epidemiologic” of the ELISA or WB, data which often remains ‘black-boxed’ from a critical public scrutiny. Using data from tested subjects and published accounts/texts, the paper deconstructs the classification of ‘risk’ embodied by official test algorithms and analyses how the ambiguity/uncertainty characteristic of antibody-test methodologies have sociological implications for ethical decision-making, self-identity and social movements.
1. Introduction
This paper analyses the epistemology of diagnosis for the human immunodeficiency virus (HIV). Diagnosis of HIV infection requires two parallel and interlocking processes: a biochemical test that takes place in a laboratory using a commercially prepared test-kit and a dialogical assessment of ‘risk of exposure’ that takes place in the clinic between individuals and health professionals.

In the first part of the paper, connections are critically explored between the laboratory identification of HIV and the clinic-based risk assessment and risk categorisation of the individual who donates their blood for HIV testing. The aim is to analyse how clinical categorisation of risk category may provide the defining context for the interpretation of the ‘objective’ signal from the biochemical reaction in the laboratory test. Drawing on algorithms published since 1986 by the United Kingdom’s (UK) Communicable Disease Surveillance Centre, I will analyse whether each process - the laboratory-based test versus the dialogical pre-test interview - is of equal importance in relation to the diagnostic outcome or whether bias and discrimination are implied.

In the second part of the paper, data from tested subjects, published accounts and other texts will be used to analyse how the classification of ‘risk’, embodied by official HIV test algorithms, demonstrates ambiguity and uncertainty, characteristic of all medical screening methodologies. Various sociological implications are explored in relation to self-identity, social movements and the critical public engagement with AIDS science and technology politically stigmatised by use of the term ‘AIDS dissidence’ or ‘AIDS denialist’. I will argue that so-called ‘AIDS dissidence’ is a misnomer used by official health authorities to hide a range of insightful caveats of such biotechnologies. Lastly, using the above analyses, I will make some closing comments about relationships between differing forms of scientific expertise and the need for a greater degree of reflexivity within AIDS science.

By ‘algorithm’ I mean the combined sequence of testing and risk categorisation that aims to both identify HIV antibodies in the laboratory from a donor’s blood sample, and assess the ‘risk category’ of an individual in a clinic. The latter is derived from sexual disclosures and/ or professionals’ deductions over ‘risk’ and the ‘risk category’ of the individual for HIV transmission. These processes jointly contribute towards the formulation of diagnosis imparted to the individual by a qualified physician. By ‘black box’, I mean the manner whereby the uncertainties contingencies and controversies underpinning these algorithms are collapsed into scientific ‘fact’ through expert assimilation.

2. The Laboratory and the Clinic
Since 1985, official health authorities, like the UK Communicable Disease Surveillance Centre at Colindale, have given leadership in developing test algorithms.

1 In the UK a statutory instrument under the Health and Medicines Act 1998 directs that all HIV testing kits supplied in UK must be accompanied by a warning that at least one confirmatory test should be undertaken following a positive test result. See, Her Majesty’s Stationery Office (1992) Statutory Instruments No.460 Public Health, England and Wales. Public Health, Scotland. The HIV Testing Kits and Service Regulations 1992. London: Her Majesty’s Stationery Office.
Laboratory-based tests, or ‘assays’, in the form of commercially available test-kits have been licensed in the UK since 1985 by the UK’s Medical Devices Agency. Globally, the two commonly used laboratory tests are the Enzyme-Linked Immunosorbent Assay (ELISA) and the Western Blot (WB). The ELISA and WB are surrogate technologies that aim to detect, not the genome of HIV, but antibodies to proteins which epidemiological studies suggest appear in the blood of those with AIDS-related illnesses and which may react with an assortment of genetically engineered monoclonal proteins in the test-kits thought of as a unique to HIV\(^3,4,5\). EIA/ WB kits are laboratory processes developed by the pharmaceutical industrial complex\(^6\), a form of biotechnology\(^7,8\), and another example of the “…new biological techniques found commercial applications during the 1970s and 1980s”\(^9\). The ELISA/WB were originally designed to ensure the safety of the blood supply and were both developed from the early 1980’s biochemical laboratory procedures used to identify HIV (then known as HTLV-III)\(^10\).

Since the start of the AIDS era, the CDSC (now part of the U.K.’s Health Protection Agency) developed algorithms for antibody-tests that advocated using a combination of sensitive and specific ELISA’s. Prior to coming onto the market, these test kits were then duly evaluated by CDSC scientists.

---

\(^4\) Unlike WB, ELISA does not separate the proteins considered evidence of an immune response to HIV. ELISA yields a yes- or no- answer; WB requires operator interpretation to identify the presence of antibodies thought to be specific to HIV proteins.

\(^5\) The sensitivity, specificity and positive predictive values of these tests are calculated based on their prior stochastic value- their statistical correlation with illness- using surrogate marker evidence of immune suppression like T-cell counts and retroviral; activity like detecting reverse transcriptase the enzyme thought characteristic of HIV.


---

In one of the earliest British publications on this topic - a book chapter reporting on the first UK AIDS Conference in Newcastle 1986 - Philip Mortimer (now a Director of the HPA) described the 1986 quality monitoring of the then available commercial tests-kits coming onto the market. He talks of that which “…exemplifies what we are looking for in an effective test”, namely:

“(in) Figure 5.11 the blood donor group are all segregated on the left, whereas many individuals in the high-risk groups give a strong signal in the test, representing positive result, and lie on the right side of the histogram. Between these two zones is a wide area in which no specimen from any group gives a signal. There is thus very good discrimination between a population of positive specimens and population made up of negative specimens.”

Later in this 1986 account Mortimer admits that the blood donors in this evaluation were “all presumed to be seronegative”. Thus in 1986 what exemplified a good test was it’s ability to discern, not just populations of positive versus negative specimens, but also positive populations from those presumed to be ‘not at risk. In this histogram\(^{11}\), false positives are shown, so revealing how the technology at the very start of the AIDS era, embodied caveats over false-positivity\(^{12}\).

Mortimer then goes on to contradict his statement by saying, “…some sera give anomalous and false positive results in many commercial assays. There are several reasons for this, but basically they all involve abuse of the specimen” (p.41) referring to how storage and handling of the specimen, not the technology itself, produces false-positives. He does not describe the reasons, but it was already then known that these tests produce ‘biologically false-positive’ and ‘inderminate’ or ‘seroequivocal’ results.

Here it is useful to remember that the interpretative flexibility and these tests, for although there maybe ‘closure’ and ‘stabilisation’ over the meaning of these tests it depends which interests one is examining. U.S. test manufactures have always published information on the interpretative flexibility of these tests, which appear less than stabilised within the manufacture’s package inserts. For example, in the package insert in one of the Abbott Laboratories (1997) test kit, it states

“… ELISA was designed to be extremely sensitive. As a result, non-specific reactions may be seen in samples from some people who, due to prior pregnancy, blood transfusion, or other exposures, have antibodies to the human cells or media in which the HIV –1 is grown for manufacture of the ELISA… in most settings it is appropriate to investigate repeatably reactive specimens by additional more specific or supplemental tests.”\(^{13}\)

---

\(^{11}\) Mortimer, P. (1987) *Investigation: the work of the Laboratories in Proceedings of the AIDS Conference 1986* Newcastle upon Tyne, UK (Jones P.ed) Ponteland, Newcastle upon Tyne: Intercept (p.41 Figure 5.1).


\(^{13}\) Abbott Laboratories (1997) *Human Immunodeficiency virus type 1. Immunoassay for the detection of antibody to human immunodeficiency virus type 1 (HIV-1) in human serum or plasma.* Abbott Park, Illinois USA: Abbott Laboratories, Diagnostic Division. P.1
Health authorities try to minimise these biases. In England and Wales repeatably reactive specimens are re-tested within algorithms combining ELISAs of differing specificity and sensitivity. In the USA, ELISAs are confirmed by Western blot tests (considered by US authorities as the most specific test for HIV antibodies). Yet Organon Teknika Corporation - one manufacturer of U.S. Western blot kits - state in their package insert:

“Clinical samples have also described that are reactive in the screening assays but do not contain HIV-1 antibody. Some of these samples possess antibody to certain class II histo-compatibility antigens that are found in some cell lines used to produce the virus. Other persons, who have had no known exposure to HIV-1, produce reactive results in the screening test for still unknown reasons. Such non-specific results are found commonly when screening tests are used in large populations. Since the psychosocial and medical implications of a positive antibody test may be devastating, it has been recommended that additional testing be performed on such samples to validate the presence of antibody specific to HIV-1.

Although a positive result may indicate infection the HIV-1 virus, a diagnosis of Acquired Immunodeficiency Syndrome (AIDS) can be made only if an individual meets the case definition of AIDS established by the Centres for Disease Control.

Do not use this kit as the sole basis of diagnosis of HIV-1 infection”
(Organon Teknika 1997 emphasis added)

This test-kit manufacturer also makes the following statement about the WB test, considered the most specific for HIV antibodies:

“Since reactivity of any degree with any virus-specific proteins present on the strip results in an Intermediate result, all samples interpreted as Intermediate should be repeated using an original specimen. In addition, individuals with indeterminate results should be followed for up to six months” (Organon Teknika 1997)

Another caveat is that these test kits don’t use the virus itself as a ‘gold standard’ to validate the test, that is, there is “no criterion taken as sufficient evidence that e.g.

---


disease is present and against which other tests can be measured”\textsuperscript{17}. In relation to this fact, leading U.S. virologist Blattner states:

“… one difficulty in assessing the specificity and the sensitivity of human retrovirus assays [cf. HIV tests] is the absence of a final gold standard. In the absence of final gold standards for HIV-1, the true sensitivity and specificity for detection of viral antibodies remains imprecise”\textsuperscript{18}.

The following statement, and similar wording, is found in U.S. package inserts for the tests published by manufactures:

“At present there is no recognised standard for establishing the presence or absence of antibodies to HIV-1 and HIV-2 in human blood”\textsuperscript{19}

The global variation is diagnosis of HIV, lack of any ‘final gold standard’ for test authentication and the already well publicised problems over sensitivity/ specificity of ELISA/WB, are cited by critical scientists as evidence for alternative technological frames of meaning for these technologies as such critics argue that the only gold standard for HIV antibody-tests must be formulated in relation to the retrovirus itself and not its associated surrogate markers. Thus, the sensitivity \textsuperscript{20} and specificity \textsuperscript{21} of such tests are known to be unknown, whilst the isolation of HIV is also cited as problematic\textsuperscript{22}, thus Mortimer has stated that current diagnosis of HIV is known to be:

“…based almost entirely on detection of antibodies to HIV, but there can be misleading cross-reactions between HIV-1 antigens and antibodies formed against other antigens, and these may lead to false-positive reactions. Thus, \textbf{it may be possible to relate an antibody response specifically to HIV-1 infection}”\textsuperscript{23}.

\textsuperscript{20} “How often the test is positive when you already know what you are testing for is present” Griner, P.F., Mayewski, R.J., Mushlin, A. (1981) Selection and interpretation of diagnostic test and procedures. Annals of Internal Medicine 94 (2):559-563.
\textsuperscript{21} How often does the test read positive when what you are testing for is known to be absent” Griner et al. (1981) op.cit.
\textsuperscript{22} “… the isolation… [of HIV] involves co-cultivation of host lymphocytes with uninfected (eg. Umbilical cord) lymphocytes in the presence of interleukin II. Virus multiplication in culture can be detected by the reverse transcription assay…” Mortimer, P. (1988) The AIDS virus and the HIV test. Medicine International 56:2334-2339
\textsuperscript{23} Mortimer’s statements also indicate that by term ‘isolation’ is meant ‘identifying’ Reverse Transcriptase, an enzyme thought to be characteristic of HIV Mortimer (1988) op.cit. p.2336.
In a manufactures package insert for a test kit of the most specific HIV test, the Western blot, a fatal caveat is similarly described:

“… persons who have had no known exposure to HIV-1, produce reactive results in the screening test for still unknown reasons … it is recommended additional testing be performed.”; a positive WB result “may indicate with the HIV-1 virus.” (emphasis added)24

Thus as the package inserts show, abuse of the specimen (as Mortimer put it in 1986), is not the only reason for false-positive results; as these are an inherent aspect of the screening technology itself, and are associated with how the assays perform in relation to the blood samples from high versus low prevalence populations (known as the Positive Prediction Value (PPV)) as well as due to cross-reactions caused by impurities in the antigen preparation and cross-reactive antibodies in the donor’s blood sample25.

The lack of official acknowledgement of this information is itself a form of ‘black-boxing’ because although technical caveats are published by test manufacturers (not always in the UK but most especially in the US, perhaps for reasons of litigation), such information may not be very accessible to test subjects and/or blood donors.26

Returning to the 1986 chapter, Mortimer described the basis of today’s test algorithms (in England and Wales):

Table 5.5 shows how the confirmatory procedure operates. A primary test is carried out in the hospital laboratory or in a transfusion laboratory, and, if a positive result is obtained, our advice is that it should be checked from another specimen. We advise both that the test is repeated and also that the specimen is referred to one of the confirmatory laboratories. If, on the other hand, a negative result is obtained, we ask a number of questions. First, is it an expected and entirely negative signal? If so, it seems reasonable to report a negative result. If on the other hand, it gives a borderline result coming fairly close to the cut-off point in the assay, or if the result is not the one that might be expected, the procedure for a positive result should be followed.” (Mortimer 1986 p.47, emphasis added)

---

26 When this debate over the validity/reliability of HIV test entered a professional journal in August/September 1999, Mortimer et al.’s rebuttal chose not to comment on these sorts of disclosures by test manufacturers. Two authors wrote an article arguing against the mandatory screening of all pregnant women and directly citing at length one HIV test manufacturers’ caveats, see Harrison, R. Corbett, K.(1999) Screening of pregnant women for HIV: the case against. The Practising Midwife. 2 (7) 24-29. Mortimer et al published a reply attempting to discredit the arguments whilst ignoring the published information from HIV test manufacturer’s package inserts. See Nicoll et al (1999) Pregnant women and testing for HIV. The Practising Midwife 2, 8, 34-37. See also the original authors’ reply. Harrison , R. Corbett, K. (1999) Author’s reply. The Practising Midwife. 2 (9) 34-35.
Thus although the test is undertaken in a laboratory with its implication of ‘objectivity’, it is still something that can be ‘expected’ because the sample is accompanied by a clinical form declaring the ‘risk group’ of the donor. If the laboratory signal is contrary to that which is ‘expected’, the sample should go back for re-testing (in the diagram, “the procedure for a positive result”): thus the context of the procedure does not appear to be ‘objective’, as implied by the notion of the laboratory test, which is contextualized within presumptions over positivity and ‘risk categorgy’.  

This 1986 test algorithm was later re-worked by Mortimer et al. in a 1992 article from the UK Public Health Laboratory services AIDS Diagnosis Working Group. The 1992 algorithm describes the risk categorisation of the blood donor (‘high/low risk’) as well as a balance of the sensitivity and specificity of the available assays. The article talks of ruling out false reactions in the initial assay A, using further screening assays B, C. It then goes on to discuss the ‘common outcomes’:

 “…either the two further assays are both unreactive or (see below) they are both reactive. The outcomes A+ B- C- is highly suggestive of false positivity in assay A, especially if the reaction in A is unrepeatable. If the optical density/cut off ration for the assay is <2 and the individual is not stated on the clinician’s request form to be at high risk, it is recommended that a negative report be issued without follow up.” (p.61, emphasis added).

The aim is to require blood samples that test positive in England and Wales to be retested with the same ELISA methodology to confirm a ‘positive’ a ‘true positive’. There is mention of the ‘high/low risk’ donor below the section of the algorithm labelled “OD/CO<2” with two options dependent upon whether the donor is considered and/or categorized as ‘low/high risk’. The paper doesn’t talk of the ‘expected’ result. But if the sample is thought to come from a ‘high risk’ person (even though it gives the same optical density signal as a ‘low’ risk sample) it is reported “Seek Follow Up Sample”, and accordingly another sample is sought for further testing.

My point here is that the same laboratory signal means something different in context of the perception of risk posed by the donor of the blood. If it’s thought the sample is from a ‘low risk’ person (even with the OD reactivity), the sample exits the algorithm with a ‘negative’ laboratory report. If the sample is perceived to come from ‘individuals within high risk groups’ the laboratory scrutiny is greater, as is the suspicion of its positive nature. Unlike the 1986 chapter the discourse of expectation

---

27 In this 1986 chapter there is no mention of how ‘risk’ is assessed in the clinic. This omission, and the collapse of the notions of ‘risk behaviour’ into ‘risk category’ together with explicit statements on ‘presumptive risk’ that underpin the algorithm, are all features of this early 1986 discourse. In England and Wales this sort of algorithm uses ELISA only, replacing the ELISA plus confirmatory Western blot methodology used by health authorities in the United States and Scotland. See, Mortimer, P (1991) The fallibility of the HIV Western Blot. *Lancet* 337 p286-287, Feb 2.

28 However, the same test methodologies are used as their own controls and as a means of confirming the initial findings with no independent measurement of the specific outcome variable, HIV. Negative results are not routinely retested or referred for confirmatory testing, neither in the UK nor in all parts of the world further adding to regional variations in diagnosis, see Mortimer, P (1992) Towards error free diagnosis. *Public Health Laboratory Service Microbiology Digest* 9 (2), 61-64.; Papadopulos-Eleopulos et al (1993) op cit, especially pp.697-698.
and presumption is not explicit, but the potential for discrimination between the samples is evident within the algorithm, and the text, which makes explicit reference to risk category in the clinician’s report:

“If the optical density/cut-off ratio for the assay A is <2 and the individual is not stated on the clinicians request from to be at high risk, it is recommended that a negative report be issued without follow up.”

In the 2003 re-worked and published version of the algorithm the above issues are not explicitly discussed in the text, or alluded to in the algorithm, but are collapsed into the following statement appearing in the text:

“On going audit of the output of any confirmatory algorithm and checking that the final result is not at odds with patients’ clinical and behavioural characteristics is a key element” (Parry et al 2003, emphasis added).

If the donor is perceived as being at ‘high-risk’ there is not only greater scrutiny of the sample but a lingering suspicion of positivity so eroding the laboratory-based test being purely ‘objective’. Given the constant of exposure category and PPV, how assays (tests) perform in differing populations, as Mariah Mensah argues (Mensah 2000), and I have demonstrated above, exposure category is one constant upon which seropositivity is measured. I further argue that the interpretation of the test ‘result’, thus the diagnosis, is not wholly ‘objective’ or laboratory-determined, as it relies as much upon the classification of the test subject (blood donor) as being ‘at risk’ during pre-test dialogue, as it does upon the “epidemo-logic” (Mensah 2000) of the ELISA or WB.


In the second part of the paper, I’m focusing on what it can mean when individuals learn of the caveats of these technologies, not from an academic perspective, but from an experiential one. For example, whilst interviewing for my PhD these, I [KEVIN] recorded the following dialogue with one interviewee (study respondent)[JAMIE] about testing:

KEVIN: How could it be that they [blood samples] are labelled like that? Do you know that they are?

JAMIE: Thousands of them. I am certain that they are.

KEVIN: How are you certain?

JAMIE: Because I have been in hospitals and I have actually read the clinician’s request forms for myself and seen the clinician’s request form. I have also got

photocopies of the clinician’s request form from various hospitals. So, I know that all good specimens are accompanied by information about the individual. This information will include whether or not you’ve had a previous positive result, what type of person you are, what type of people you have had sex with, whether or not you have been at risk of so-called HIV in the last six months. So it includes all that information that goes along with your blood sample to the laboratory where the test is performed. Now for me this is not on. The first answer to the question, the first answer that you know is that this information, the clinician’s request form which accompanies the blood sample to the laboratory is used only for epidemiological information and for the public health laboratory to know what the infection rate is like in those groups. I think that is bullshit myself because the doctors could report any positive results and I think that given the claims made by the doctors about the accuracy, the statistically, the reproducibility and the reliability of these antibody tests, why would you need to label the blood sample? Why can’t the laboratory decide he’s negative and he’s positive, without knowing what type of person that the specimen came from? To me it is absolutely essential that all the blood samples are treated objectively and are treated and interpreted in the same way. You know, the same results should be interpreted the same way. To me it is essential that, and I feel very strongly about it that. Laboratory technicians have to test these blood samples completely by. No information about the individual at all should ever be sent, under any conditions, to a laboratory. We have to decide whether or not the test is supposed to decide, whether or not the test is supposed to tell them whether a person is positive, not additional information. If the test can’t tell you that then you should find out how often the test can’t tell you that. So really the system at present is hiding and masking all of the inconsistencies with the HIV antibody test. I think it is the inconsistencies that we should be looking at. I think we could learn a lot from the inconsistencies of the antibody test and the other tests that we are currently using. So there is that, one reason and I am actually concerned...” (emphasis added)

This sort of interpretation of the testing may be difficult for healthcare professionals to appreciate especially if working from a ‘deficit model’ of public understanding of science, and if such responses engage with so-called ‘dissenting’ opinions on HIV and AIDS.

3. Conclusions

The interpretation of indeterminate results is problematic given the possibility of false results, which are ruled out by estimating the risk of exposure (‘window period’) and the seroprevalence in the population of the test subject. Based upon the preceding analysis, the interpretation of the test ‘result’ is not wholly ‘objective’, or laboratory-determined, as it relies as much upon the classification of the test subject as being ‘at risk’ during pre-test dialogue as it does upon the “epidemi logic” of the ELISA or WB. The blood samples that test positive are re-tested (using a freshly drawn sample) with the same test methodology and with differing commercially available test-kits, in order to confirm a ‘positive’ a ‘true positive’. In this way the same test methodologies are used as their own controls and as a means of confirming the initial findings. There is no independent measurement of the specified outcome variable, HIV. Negative results are not routinely re-tested or referred for confirmatory testing in all parts of the

world. This further adds to regional variation in definitive diagnosis\(^\text{14}\). Globally, public health officials publish differing algorithms for testing of human subjects.