

# Stochastic processes defining sensitivity and variability of internally calibrated quantitative NASBA-based viral load assays

Jos J. A. M. Weusten, Pieter A. W. M. Wouters\*, Martien C. A. van Zuijlen<sup>2</sup> and Paul A. van de Wiel<sup>1</sup>

Departments of Biomathematics and <sup>1</sup>Department of Molecular Diagnostics, BioMérieux, Boxtel, The Netherlands and <sup>2</sup>Department of Mathematics, Catholic University of Nijmegen, Nijmegen, The Netherlands

Received July 11, 2002; Revised and Accepted October 7, 2002

## ABSTRACT

**For quantitative assessment of virus particles in patient plasma samples various assays are commercially available. Typical performance characteristics for such assays are sensitivity, precision and the range of linearity. In order to assess these properties it is common practice to divide the range of inputs into subranges in order to apply different statistical models to evaluate these properties separately. We developed a general statistical model for internally calibrated amplification based viral load assays that combines these statistical properties in one powerful analysis. Based on the model an unambiguous definition of the lower limit of the linear range can be given. The proposed method of analysis was illustrated by a successful application to data generated by the NucliSens EasyQ HIV-1 assay.**

## INTRODUCTION

Viral load assays based on amplification technology are used all over the world for screening and monitoring purposes. Proficiency studies and quality assurance programs are tools to monitor the quality of the assay systems themselves as well as of the laboratory staff that use them in practice. Important items include assay variability, detectability and the lower limit of the linear range. Typically, in assay design and development as well as in proficiency testing, data obtained by testing a series of concentrations are analyzed using a variety of statistical techniques to assess these properties (see for example 1). These techniques include different sorts of analysis of variance models to assess quantitative properties and probit or logistic regression to assess sensitivity. In order to apply these models, the dynamic range tested has to be divided (more or less arbitrarily) into subgroups to ensure that the heterogeneity of variances within these subgroups is not too great, whereas it is known that the assay variability, defined as the standard deviation of the logarithm of the

quantitation result, depends on the viral load tested. It would be beneficial to develop a single mathematical/statistical model that covers all properties. The current paper aims to develop such a model. Although many details of the current studies are applicable to both NASBA-based and PCR-based viral load assays, the focus is on the former.

One of the commercially available internally calibrated NASBA-based viral load monitoring systems is the NucliSens product line. The list of applications is long and includes both RNA and DNA targets (see for example 2). Important targets in the quantitative field are HIV-1 RNA (3–6) and HBV DNA (7). The NASBA amplification technology was successfully combined with homogeneous detection using molecular beacons (2,6,8–10). These oligonucleotides are designed such that they can generate a fluorescence signal only when they are bound to a nucleic acid strand with a complementary sequence, in this case an amplicon. As the concentrations of amplicons increase continuously during the NASBA process, the time-dependency of the fluorescence signals provides information on the underlying kinetics of the NASBA-driven RNA formation (6,8–10).

The NucliSens product line is, amongst others, characterized by the feature of internal calibration. A fixed amount of artificial RNA or DNA, referred to as the calibrator, is added to the sample prior to the assay. This calibrator is co-extracted and co-amplified together with the endogenous viral nucleic acid. The calibrator differs from the endogenous nucleic acid by a short sequence. The use of specific molecular beacons with specific fluorophores allows simultaneous specific detection of both amplification processes. Quantitation is based on the assessment of the relative RNA growth rates during the transcriptional phase of the amplification (10).

In a previous paper we presented the results of modeling studies on the relation between the input concentrations and transcription rates (10) and we presented a mathematical model describing the fluorescence curves as a function of time and several relevant parameters. In the current paper we wish to go a step further and to include stochastic elements that play an important role in molecular diagnostics. As such, a general mathematical model is obtained that describes several performance characteristics in a single model. The procedures

\*To whom correspondence should be addressed at BioMérieux, PO Box 84, 5280 AB Boxtel, The Netherlands. Tel: +31 411 65 46 25; Fax: +31 411 65 43 11; Email: pieter.wouters@eu.biomerieux.com

will be illustrated using data generated with the NucliSens EasyQ HIV-1 assay. This assay is also used as leading example in the theoretical sections.

## THEORY

### Overview of the assay and some definitions

Consider a patient with a given load of endogenous viral RNA [referred to as wild-type (WT) RNA]. From this patient a blood plasma sample is taken. Let the expectation of the number of WT copies in this particular volume of plasma be  $N$ . For a random sample drawn from the patient, it should be realized that the actual number of copies present is a random variable. Let the plasma sample contain  $N_1$  copies. Then calibrator RNA is added to this patient sample. In the NucliSens assays, this calibrator is added in dry form as a so-called accuser. These accusers are produced by freeze-drying small aliquots from a large pool of a solution of the calibrator RNA. The concentration of calibrator RNA in this pool is such that, on average, a single accuser contains  $C$  copies. Similar as for the patient sample, the actual number of copies in a single accuser is a random variable. Let an individual accuser contain  $C_1$  copies. Both WT and calibrator RNA are simultaneously subjected to the extraction. The numbers of copies present in the extract are  $N_2$  and  $C_2$ . From this extract a fraction is used for the amplification, containing  $N_3$  and  $C_3$  copies, respectively. These copies enter the amplification, but not all copies are amplified to yield a detectable number of amplicons. Let the number of copies that actually contribute to the total signal be  $N_4$  and  $C_4$ . The response to be obtained in the NASBA-based assay is determined by the ratio of the amplified copy numbers ( $N_4 / C_4$ ), and is expressed on a logarithmic scale (base 10). More precisely, the response variable obtained reflects the logarithm of the transcription rate ratio, which is proportional to the input ratio  $N_4 / C_4$  (10). The observed value is converted to a viral load (to be referred to as  $N_{\text{obs}}$ ) by using the so-called batch parameters. Obviously, random error due to the NASBA assay is added to the total variation of the response variable. Figure 1 gives a schematic representation of all individual steps.

### Model building

For the construction of a mathematical model, the following considerations and assumptions are relevant. (i) The values of  $N$  and  $C$  are known and fixed. (ii) The variables  $N_1, N_2, N_3, N_4, C_1, C_2, C_3, C_4, \varepsilon, R$  and  $N_{\text{obs}}$  are random variables. (iii) The variables  $N_1$  and  $C_1$  follow Poisson distributions. (iv) Given the realizations of  $N_1$  and  $C_1$ , the values  $N_2$  and  $C_2$  follow binomial distributions, with a  $p$ -value determined by the extraction efficiency (to be referred to as  $r$ ). (v) Given the realizations of  $N_2$  and  $C_2$ , the values  $N_3$  and  $C_3$  follow binomial distributions, with a  $p$ -value equal to the fraction taken (to be referred to as  $f$ ). (vi) Given the realizations of  $N_3$  and  $C_3$ , the values  $N_4$  and  $C_4$  follow binomial distributions, with a  $p$ -value equal to the probability that an individual molecule is amplified to yield a detectable number of amplicons (to be referred to as  $p$ ). (vii) The value of  $\log(N_4 / C_4)$  is to be assessed by the NASBA-based assay. (viii) The NASBA assay itself (amplification and detection) induces a

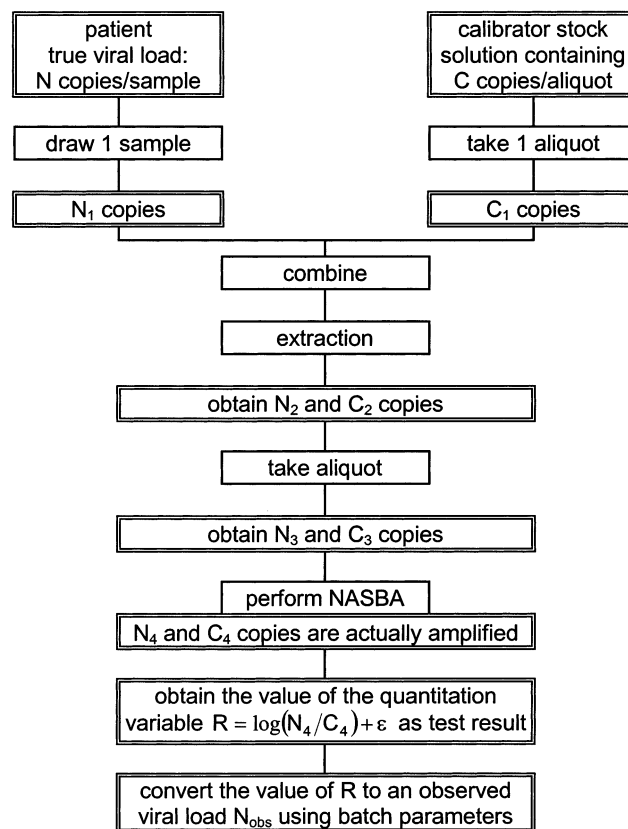


Figure 1. Schematic overview over the model.

random variation  $\varepsilon$  that is normally distributed with expectation 0 and variance  $\sigma_a^2$ . (ix) The quantity  $\log(N_4 / C_4) + \varepsilon$  is the final response variable, to be referred to as  $R$ . (x) The conversion of  $R$  to  $\log(N_{\text{obs}})$  is done by a linear transformation using a fixed set of numbers referred to as the batch parameters. A test result is classified as positive if two conditions are fulfilled simultaneously: (i) a WT signal is detected and (ii) the observed value of  $N_{\text{obs}}$  exceeds the threshold  $N_{\text{lim}}$ . In the other cases the estimated viral load is below the limit of detection ('negative').

### Probability distributions of the relevant random variables

The variables  $N_1$  and  $C_1$  are considered to follow Poisson distributions as it seems reasonable to assume that the three basic probabilistic conditions for Poisson variables are fulfilled. First, the expectation of the number of molecules in the sample is proportional to the sample volume. Second, two subsequent samplings from the same pool are stochastically independent and identically distributed. Third, if the volume of the sample is small enough, the probability that two molecules are present in the same sample approaches 0. Therefore,  $N_1 \sim \text{Poi}(N)$  and  $C_1 \sim \text{Poi}(C)$ , with the symbol  $\sim$  to indicate a probability distribution.

As summarized above, given the realization of the random variable  $N_1$ , the variable  $N_2$  follows a binomial probability distribution with parameter  $r$ . General stochastic theory then shows that  $N_2 \sim \text{Poi}(rN)$ . Similarly, it follows for the other steps in the scheme:

$$\left\{ \begin{array}{ll} \text{copies in blood sample or accusphere :} & N_1 \sim \text{Poi}(N), \quad C_1 \sim \text{Poi}(C) \\ \text{copies in extract :} & N_2 \sim \text{Poi}(fN), \quad C_2 \sim \text{Poi}(fC) \\ \text{copies in fraction } f \text{ of eluate :} & N_3 \sim \text{Poi}(fprN), \quad C_3 \sim \text{Poi}(fprC) \\ \text{amplified copies :} & N_4 \sim \text{Poi}(fprN), \quad C_4 \sim \text{Poi}(fprC) \end{array} \right. \quad \mathbf{1}$$

**The response variable**

The response variable  $R$  is defined as

$$R = \log(N_4 / C_4) + \varepsilon \text{ with } \varepsilon \sim N(0, \sigma_a^2) \quad \mathbf{2}$$

with  $\log$  the logarithm with base 10 and  $N$  reflecting a normal distribution. Note that  $R$  is defined only if the realizations of  $N_4$  and  $C_4$  both exceed 0. This notion has some important implications, as will be shown below. The total variability in this variable is defined by the variability in  $N_4$ ,  $C_4$  and  $\varepsilon$ . The observed value of  $R$  is converted to a logarithmic viral load  $[\log(N_{\text{obs}})]$  by a linear transformation (10). To this end, batch parameters are used. For reasons of convenience, this transformation is written as

$$\log(N_{\text{obs}}) = q_1 + q_2 R \quad \mathbf{3}$$

with  $q_1$  and  $q_2$  reflecting the batch parameters (the exact formal definition of the batch parameters is slightly different, but here only the linear relation is of importance). As described previously (10), it is not possible to directly assess the ratio  $N_4 / C_4$ ; the quantitation variable merely reflects the logarithm of some multiple of the  $N_4 / C_4$  ratio, but batch parameters can be used to convert this to an observed viral load.

The batch parameters are assessed once and are then kept fixed. Therefore, in discussions on random errors they are not important.

**Probabilistic aspects of the response variable**

The variable  $R$  is the response variable used for quantitation. If the realization of  $N_4 = 0$ , then a test result without detectable WT signal is obtained, which is classified as ‘negative’. In this case  $R$  is not computed at all. If there is a detectable WT response, it means that the realization of  $N_4 > 0$  and in that case the value of  $R$  can be computed. If the estimated viral load is below the lower limit of detection, the result is classified as ‘negative’, otherwise it is classified as ‘positive’ and a quantitation result is presented. In order to assess the variance of  $R$ , use has to be made of the probability distribution of the logarithm of  $N_4$  under the restriction that the realization of  $N_4 > 0$ . Some details of the properties of the logarithm of such a truncated Poisson variable can be found in the Appendix (equations **A1** and **A2**).

A realization of 0 for the variable  $C_4$  implies the absence of a calibrator signal. Therefore, similar considerations are required for the probability distribution of  $C_4$  as described for  $N_4$ . However, as the value of Poisson parameter  $fprC$  (see equation **1**) is considered to be fixed for a given batch, there is no need to use explicit equations for the expectation and variance of  $\log(C_4)$ , as will become clear from the derivations presented below.

From equations **2** and **3** it follows that the random variable  $\log(N_{\text{obs}})$  can be written as the sum of three independent random variables:  $\log(N_4)$ ,  $\log(C_4)$  and  $\varepsilon$ . Therefore,

$$\begin{aligned} E[\log(N_{\text{obs}})] &= q_1 + q_2 \{E[\log(N_4)] - E[\log(C_4)] + E[\varepsilon]\} \\ S_{[\log(N_{\text{obs}})]}^2 &= q_2^2 S_{[R]}^2 = q_2^2 \{S_{[\log(N_4)]}^2 + S_{[\log(C_4)]}^2 + S_{[\varepsilon]}^2\} \end{aligned}$$

with  $E[.]$  the expectation and  $S_{[.]}^2$  the variance of the random variable indicated. In order to write these in manageable forms, define

$$\left\{ \begin{array}{l} \gamma_1 = fpr \\ \gamma_2 = q_1 + q_2 \log(fpr) - q_2 E[\log(C_4)] \\ \gamma_3 = q_2 \\ \gamma_4 = q_2^2 \{S_{[\log(C_4)]}^2 + \sigma_a^2\} \end{array} \right. \quad \mathbf{4}$$

and it follows using the Appendix equations

$$E[\log(N_{\text{obs}})] = \gamma_2 + \gamma_3 \{ \log(N) + \{[(\gamma_1 N + 1)e^{-\gamma_1 N} - 1] / [2\gamma_1 N \ln(10)]\} - \log(1 - e^{-\gamma_1 N}) \} \quad \mathbf{5}$$

$$S_{[\log(N_{\text{obs}})]}^2 = (\gamma_3 / \ln 10)^2 / [\gamma_1 N - \vartheta_1 + \vartheta_2 / (\gamma_1 N)^{\vartheta_3}] + \gamma_4 \quad \mathbf{6}$$

As noted above, it is possible to express  $E[\log(C_4)]$  and  $S_{[\log(C_4)]}^2$  in terms of  $fpr$  and  $C$  using the Appendix equations, but given a value for  $fprC$  these two are constants and are incorporated in  $\gamma_2$  and  $\gamma_4$ . Nevertheless, it is important to note that  $C$  is so high that  $\text{Pr}(C_4 = 0)$  can be ignored. In other words,  $e^{-fprC}$  can be set to 0 (see equation **1**) and  $E[\log(C_4)]$  can be approximated by  $\log(fprC)$ . Therefore, it follows

$$\gamma_2 = q_1 - q_2 \log(C) \quad \mathbf{7}$$

independent of the product  $fpr$ .

Numerical analysis of equation **5** shows that, in approximation,

$$E[\log(N_{\text{obs}})] = \gamma_2 + \gamma_3 \log(N) \text{ if } \gamma_1 N = 1.9797 \quad \mathbf{8}$$

This result is important in defining the lower limit of the linear range of the assay, and will be rounded to  $\gamma_1 N = 2$ .

It is interesting to note the following limits that can be derived:

$$\left\{ \begin{array}{l} \lim_{N \rightarrow 0} E[\log(N_{\text{obs}})] = \gamma_2 - \gamma_3 \log(\gamma_1) \\ \lim_{N \rightarrow \infty} \{E[\log(N_{\text{obs}})] - [\gamma_2 + \gamma_3 \log(N)]\} = 0 \\ \lim_{N \rightarrow \infty} S_{[\log(N_{\text{obs}})]}^2 = \lim_{N \rightarrow 0} S_{[\log(N_{\text{obs}})]}^2 = \gamma_4 \end{array} \right. \quad \mathbf{9}$$

Equation **5** shows that the relationship between the input and the output, that is between  $\log(N)$  and  $E[\log(N_{\text{obs}})]$ , is non-linear in nature. For high values of  $N$ , equation **9** shows that  $E[\log(N_{\text{obs}})]$  can be approximated by  $\gamma_2 + \gamma_3 \log(N)$ . The difference between these two is given primarily by the entity  $\log(1 - e^{-\gamma_1 N}) - [(\gamma_1 N + 1)e^{-\gamma_1 N} - 1] / [2\gamma_1 N \ln(10)]$ , which reaches its maximum of 0.041 for  $\gamma_1 N$  approximately 4. For very low inputs, with  $N$  approaching 0, equation **9** shows that the expectation of  $\log(N_{\text{obs}})$  is well-defined and it approaches a value in which  $\gamma_1$  plays an important role.

A test result is classified as ‘positive’ if two conditions are fulfilled: a WT signal is detected ( $N_4 > 0$ ) and the quantitation result exceeds a pre-set lower limit ( $N_{\text{obs}} > N_{\text{lim}}$ ). More

precisely, the probability that a result is positive is the product of the probability  $\Pr(N_4 > 0)$  and the conditional probability  $\Pr(N_{\text{obs}} > N_{\text{lim}} \mid N_4 > 0)$ . Using the Normal approximation it follows:

$$\Pr(\text{pos}) = (1 - e^{-\gamma_1 N}) [1 - \Phi(\{\log(N_{\text{lim}}) - E[\log(N_{\text{obs}})]\} / \sqrt{\{S_{[\log(N_{\text{obs}})]^2}\})}] \quad 10$$

with  $E[\log(N_{\text{obs}})]$  and  $S_{[\log(N_{\text{obs}})]^2}$  according to equations 5 and 6 and  $\Phi(\cdot)$  reflecting the cumulative standard normal probability distribution function.

### Data analysis using the current models: maximum likelihood

Consider an experiment in which samples with viral loads over a considerable range are subjected to the NucliSens EasyQ HIV-1 assay. These samples yield either a quantitation result, or are classified as below the limit of detection ('negative'). The values of the parameters given in equation 4 can be assessed by using maximum likelihood techniques. The contribution to the total log likelihood of a test result classifying as below the limit of detection at a given level of  $N$  is given by the probability that a test result is obtained that is not classified as positive, and is computed using a cumulative probability distribution function as (see equation 10):

$$L_{\text{neg}} = \ln\{1 - (1 - e^{-\gamma_1 N}) [1 - \Phi(\{\log(N_{\text{lim}}) - E[\log(N_{\text{obs}})]\} / \sqrt{\{S_{[\log(N_{\text{obs}})]^2}\})}]\} \quad 11$$

For a positive, quantifiable result, the contribution to the total log likelihood is based on the probability density function under the restriction that  $N_4 > 0$  in combination with the probability  $\Pr(N_4 > 0)$ . Let  $x_i(N)$  be the realization of  $\log(N_{\text{obs}})$  for observation  $i$  at a given value of  $N$ , and the contribution of this particular data point to the total log likelihood can be expressed as

$$L_{\text{pos}} = \ln\{(1 - e^{-\gamma_1 N}) (1 / \sqrt{\{S_{[\log(N_{\text{obs}})]^2}\}}) \phi(\{x_i(N) - E[\log(N_{\text{obs}})]\} / \sqrt{\{S_{[\log(N_{\text{obs}})]^2}\}})\} \quad 12$$

In equations 11 and 12 the entities  $E[\log(N_{\text{obs}})]$  and  $S_{[\log(N_{\text{obs}})]^2}$  are computed according to equations 5 and 6, and  $\phi(x) = [1 / \sqrt{(2\pi)}] \exp[-(1/2)x^2]$ . For each individual observed test result the contribution to the log likelihood is computed by equation 11 or 12, and all contributions are summated. Numerical techniques can then be used to assess the values of the parameters  $\gamma_1$  through  $\gamma_4$  such that the total log likelihood is maximized.

## MATERIALS AND METHODS

Dilution series of the so-called VQA stock of HIV-1 RNA (Virology Quality Assurance) (1) were subjected to the NucliSens EasyQ HIV-1 assay, the details of which are described elsewhere (6,8). All results were obtained using a single batch. For the current study magnetic extraction was used as front-end (7), in which the extract can be concentrated into a smaller volume than in the current NucliSens Extractor (11).

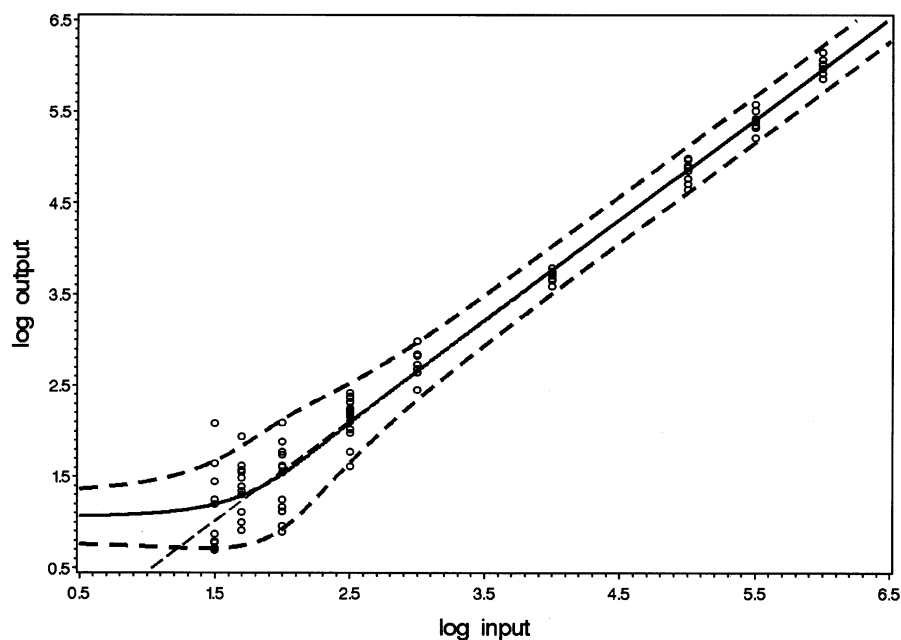
## RESULTS

Dilution series of WT HIV-1 RNA with nominal concentrations ranging from 10 to  $10^6$  copies per input were subjected to the NucliSens EasyQ HIV-1 assay. Figure 2 presents the observed quantitative results. Test results classifying as 'below limit of detection' are not presented. Using maximum likelihood techniques, the parameters of the model were assessed. For the current data, the optimal parameter values were  $\gamma_1 = 0.0295$ ,  $\gamma_2 = -0.629$ ,  $\gamma_3 = 1.10$  and  $\gamma_4 = 0.0163$ . The figure includes the predicted values of  $E[\log(N_{\text{obs}})]$  according to equation 5 as well as the intervals  $E[\log(N_{\text{obs}})] \pm 2\sqrt{\{S_{[\log(N_{\text{obs}})]^2}\}}$  (equations 5 and 6). Finally, the straight line with intercept  $\gamma_2$  and slope  $\gamma_3$  is included. According to equation 9, the expectation  $E[\log(N_{\text{obs}})]$  approaches this line for high values of  $N$ . The figure shows that this approximation is reasonable for all inputs exceeding the point where the curve of  $E[\log(N_{\text{obs}})]$  crosses this straight line. According to equation 8, this intersection is approximately at  $\gamma_1 N = 2$ , hence at  $N = 67$ . Figure 2 shows that this intersection point can be used as a definition for the lower end of the linear range. For the current data,  $\lim_{N \rightarrow 0} E[\log(N_{\text{obs}})] = 1.05$  logs.

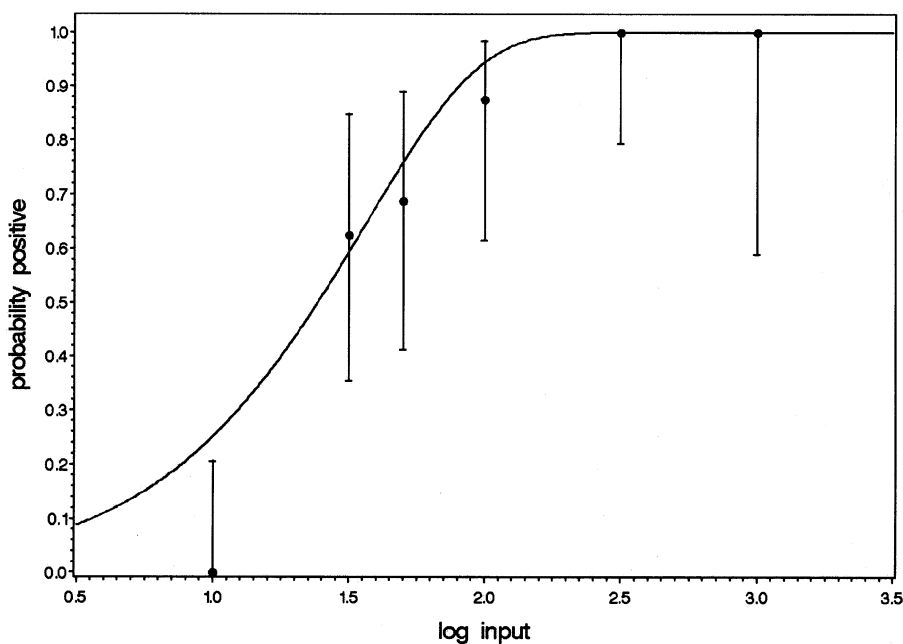
The bend in the curve corresponding to  $E[\log(N_{\text{obs}})]$  in Figure 2 is due to stochastic properties of the logarithm of a Poisson distributed variable (compare Figure 2 with Figure 4 in the Appendix). Recall that the predicted values for the expectation and variance of  $\log(N_{\text{obs}})$  are valid only for results of which a detectable WT signal is detected in the first place, and this probability rapidly decreases with decreasing values of the input. At very low inputs, with  $fprN$  values clearly  $< 1$ , it follows that the actual values that  $N_4$  can reasonably attain are only 0 and 1; the probability  $\Pr(N_4 > 1)$  is very low if  $fprN \ll 1$  (see equation 1). Therefore, at these inputs, the mere detection of a WT signal implies that the actual realization of  $N_4$  can hardly be anything else but 1. At this level of  $fprN$  values, information on the actual value of  $N$  is lost.

The model also allows predictions on the variability as a function of the input (equation 6). The assay variability is predicted to stabilize for high WT inputs at a value of  $\sqrt{\gamma_4} = 0.128$  logs. The maximal predicted variability is obtained at an input of 92 copies, and equals  $\sqrt{0.0899} = 0.30$  logs. At an input of 400 copies, the predicted standard deviation is 0.20 logs. Equation 9 (last part) shows that the predicted standard deviations for very low and very high inputs are asymptotically the same, which is also reflected in Figure 2. At very high inputs, the variability in  $\log(N_4)$  is so small compared with the variability in  $\log(C_4) + \epsilon$  that it can be ignored. At very low inputs, the probability  $\Pr(N_4 > 1)$  can be ignored, and the probability distribution of  $N_4$  under the restriction that it is positive (otherwise  $R$  is not defined) is a degenerated probability distribution with zero variance, leading to the same result.

Finally, given the parameter values, the predicted probability of obtaining a positive test result can be computed as a function of the input using equation 10. The observed and predicted hit rates are presented in Figure 3. The concentration yielding a hit rate of 95% is estimated to be 103 copies/input, whereas a 50% hit rate is obtained at an input of 24 copies. Figure 3 also includes the two-sided 95% confidence intervals that can be computed for the individual observed hit rates using the binomial probability distribution. The observed hit



**Figure 2.** Model concerning quantitative results. The circles present the individual observations. The solid line presents the modeled values of  $E[\log(N_{\text{obs}})]$ . The dashed lines are at a distance of two times the predicted SD. The intercept and slope of the straight dashed line are given by  $\gamma_2$  and  $\gamma_3$ .



**Figure 3.** Model concerning hit rates. The dots represent the observed fractions of positive test results, with the two-sided 95% confidence intervals based on the binomial probability distribution. The number of observations varied with the inputs tested.

rate at 10 copies/input is rather low compared with the predicted hit rates, but for the remaining part the model adequately describes the data.

## DISCUSSION

In the current paper the stochastic processes involved in molecular diagnostics are studied from a mathematical point of view. The study shows that a single statistical model can be

used to study linearity, variability and detectability. The lower limit of the linear range can be assessed using clear and unambiguous definitions, and is not hampered by issues like heterogeneity of variances over the linear range or low hit rates. Input concentrations corresponding to pre-set hit rates can be estimated, which is important with respect to the sensitivity of the assay. In addition, it is possible to estimate the assay variability as a function of the input, which is amongst others important in assessing the limit of quantitation.

As is the case for every model, the equations presented here are approximations of reality. For example, it is assumed that the variance component due to the assay itself ( $\sigma_a^2$ ) does not depend on the input. At very high inputs this is not entirely true, as the calibrator signal is hardly detectable and therefore the contribution of the calibrator to the variability may be higher than in other ranges. A similar notion might be valid for the low inputs, where the WT signal is on the edge of being detectable. Another approximation is the use of the Normal probability distribution to assess hit rates (equation 10) and in the maximum likelihood techniques (equations 11 and 12). Finally, the current equations assume that  $fpr$  ( $\gamma_1$ ) can be considered to be a constant, whereas it is more realistic to assume that there is a random component to this as well. This parameter makes its appearance in lots of positions in the equations, thereby making detailed mathematical analyses assuming a stochastic nature for  $fpr$  very complex. Further studies are required to expand the models to include this source of variation.

The data analyzed in the current paper were obtained using a single batch over several experimental runs. As such, the standard deviation contains both the inter- and intrarun variability. Studies have shown that the interrater variability is so small that it can safely be ignored (T. A. M. Oosterlaken and P. A. van de Wiel, unpublished results). If this were not the case, then the models should be expanded to include indicators for the experimental runs involved.

The NucliSens assay is an internally calibrated assay. Changes in extraction efficiency should not affect the quantitation. That this is actually achieved is reflected by equation 7 which shows that the location of the straight line describing the linear range of the assay is not affected by the actual value of the product  $fpr$ .

The predicted hit rate at the lowest input is clearly higher than the observed hit rate. This is probably due to the following considerations, which are very difficult to include in mathematical models. An assay should be sensitive, implying that a high hit rate at low concentrations is a beneficial property. On the other hand, the incidence of false-positive results should be low. The positive predictive value, which is the probability that the sample tested does actually contain the analyte given that the assay result is positive, should be very high. Therefore the threshold to be used to decide whether or not a WT signal is actually present is chosen higher than theoretically minimally required. As such, WT signals just above the background will be dismissed as just noise.

A test result is classified as being below the limit of detection ('negative') if no WT signal is detected or if the estimated viral load, in the presence of a detectable WT signal, is below a threshold. The equations on the variance and expectation of  $\log(N_{\text{obs}})$  describe the complete distribution of the test results under the restriction that a WT signal is observed (that is,  $N_4 > 0$ ). This means that, at low viral loads, only a part of the distribution will be visible, whereas the results below the limit of detection with a detectable signal can be treated as left-censored observations. The models show that high quantitation results cannot be obtained at very low inputs, as both  $E[\log(N_{\text{obs}})]$  and  $S_{[\log(N_{\text{obs}})]^2}$  stabilize with  $N$  approaching 0.

The product  $fpr$  ( $\gamma_1$ ) plays an important role in the equations. It makes its appearance at a crucial position in

the equation on detectability (equation 10). The factor  $1 - e^{-\gamma_1 N}$  describes the probability that at least one WT copy enters the amplification that contributes to the total signal [ $\Pr(N_4 > 0)$ ], and represents an upper bound for the probability of obtaining a positive test result as a function of the input. The other factor in equation 10 gives the probability that a positive result is obtained under the condition that a molecule enters the amplification in the first place. The lower limit of the linear range is given by  $2 / \gamma_1$  (equation 8), hence a high value of the product  $fpr$  implies a lower limit at a low input. Closely related to this is the level at which  $E[\log(N_{\text{obs}})]$  stabilizes for low inputs. Once again, a higher value of  $fpr$  ( $\gamma_1$ ) implies a further extension of the linear range to lower inputs (see equation 9). If sensitivity of the assay is to be improved or, analogously, the linear range is to be extended to lower inputs, then the efforts should focus on increasing the value of the product  $fpr$ . In more practical terms this means that  $f$  (the fraction of the eluate used),  $r$  (the extraction recovery) or  $p$  (probability that a single molecule yields a detectable signal) should be increased. The strategy behind the development of magnetic extraction is, amongst others, to increase the value of  $f$ , the fraction of the eluate used, whereas the technique is under development to further increase the recovery  $r$ . Increasing the value of  $p$  would require further optimization of the amplification module and/or the quality of the enzymes. Note that the value of  $N$  stands for the expectation of the total number of WT copies entering the assay (see Fig. 1); the sample volume is not part of the equations. As such, as has indeed been shown experimentally (12), increasing the sample volume will have a direct impact on sensitivity if this is to be expressed in terms of copies/ml but not when it is expressed as copies/input. On the other hand, the use of high plasma volumes might also lead to increased concentrations of plasma components that can interfere with the amplification, hence reducing  $p$ .

The input concentration corresponding to a hit rate of 95% is often used as a definition of the limit of detection (LOD). Interestingly, the models imply that the lower end of the linear range is by definition below this LOD. After all, according to equation 8 the lower end is at an input such that  $\gamma_1 N = 2$ . At this input equation 10 predicts that  $\Pr(\text{pos}) \leq 1 - e^{-\gamma_1 N} = 1 - e^{-2} \approx 0.86$ . The quantitative linear range is often defined as the concentration range for which there is a linear relation between the logarithm of the input and  $E[\log(N_{\text{obs}})]$  and for which the probability of obtaining a positive test result  $>95\%$ . This can further be extended to include a minimal assay standard deviation. The results of the current study indicate that with respect to the first two items, the 95% detectability requirement is sufficient.

As illustrated in the current paper, it is possible to apply the model to experimental data. For the first time, it is possible to draw conclusions on accuracy, precision and sensitivity based on a single analysis. There is no need to distinguish, on arbitrary grounds, an input range for which the variability is assumed to be constant. Also, as has been reported elsewhere (J.J.A.M. Weusten, P.A.W.M. Wouters and M.C.A. van Zuijlen, submitted for publication), effective use is made of the available information in the data set when it comes to sensitivity analyses, whereas the standard probit-analyses only use the dichotomous result 'positive' or 'negative'.

In conclusion, a mathematical model was developed that enables simultaneous analysis of linearity, variability and

detectability. It is believed that models like these can be helpful in developing and optimizing viral load diagnostics, as well as in statistical data analysis. In addition, they can contribute to the scientific knowledge of these systems. Further studies are required to refine the models.

## REFERENCES

1. Yen-Lieberman, B., Brambilla, D., Jackson, B., Bremer, J., Coombs, R., Cronin, M., Herman, S., Katzenstein, D., Leung, S., Ju Lin, H., Palumbo, P., Rasheed, S., Todd, J., Vahey, M. and Reichelderfer, P. (1996) Evaluation of a quality assurance program for quantitation of human immunodeficiency virus type 1 RNA in plasma by the AIDS clinical trials group virology laboratories. *J. Clin. Microbiol.*, **34**, 2695–2701.
2. Deiman, B., Van Aarle, P. and Sillekens, P. (2002) Characteristics and applications of Nucleic Acid Sequence-Based Amplification (NASBA). *Mol. Biotechnol.*, **20**, 163–179.
3. Van Gemen, B., Kievits, T., Nara, P., Huisman, H.G., Jurriaans, S., Goudsmit, J. and Lens, P. (1993) Qualitative and quantitative detection of HIV-1 RNA by nucleic acid sequence-based amplification. *AIDS*, **7** (suppl 2), S107–S110.
4. Van Gemen, B., Van Beuningen, R., Nabbe, A., Van Strijp, D., Jurriaans, S., Lens, P. and Kievits, T. (1994) A one-tube quantitative HIV-1 RNA NASBA nucleic acid amplification assay using electrochemiluminescent (ECL) labelled probes. *J. Virol. Methods*, **49**, 157–168.
5. Murphy, D.G., Côté, L., Fauvel, M., René, P. and Vincelette, J. (2000) Multicenter Comparison of Roche COBAS AMPLICOR MONITOR Version 1.5, Organon Teknika NucliSens QT with Extractor, and Bayer Quantiplex Version 3.0 for Quantification of Human Immunodeficiency Virus Type 1 RNA in Plasma. *J. Clin. Microbiol.*, **38**, 4034–4041.
6. Van Beuningen, R., Marras, S.A., Kramer, F.R., Oosterlaken, T., Weusten, J., Borst, G. and Van de Wiel, P. (2001) Development of a high throughput detection system for HIV-1 using real-time NASBA based on molecular beacons. In Raghavachari, R. and Tan, W. (eds), *Genomics and Proteomics Technologies, Proceedings of SPIE*. Society of Photo-Optical Instrumentation Engineers, Washington, DC, Vol. 4264, pp. 66–72.
7. Van Deursen, P., Verhoeven, A., Haarbosch, P., Verwimp, E. and Kreuvel, H. (2001) A new automated system for nucleic acid isolation in the clinical lab. *J. Microbiol. Meth.*, **47**, 124 (congress abstract).
8. Leone, G., Van Schijndel, H., Van Gemen, B., Kramer, F.R. and Schoen, C.D. (1998) Molecular beacon probes combined with amplification by NASBA enable homogeneous, real-time detection of RNA. *Nucleic Acids Res.*, **26**, 2150–2155.
9. Van de Wiel, P., Top, B., Weusten, J. and Oosterlaken, T. (2000) A new, high throughput NucliSens assay for HIV-1 viral load measurement based on real-time detection with molecular beacons. XIII International AIDS conference, 9–14 July, 2000, Durban, South Africa.
10. Weusten, J.J.A.M., Carpay, W.M., Oosterlaken, T.A.M., Van Zuijlen, M.C.A. and Van de Wiel, P.A. (2002) Principles of quantitation of viral loads using nucleic acid sequence-based amplification in combination with homogeneous detection using molecular beacons. *Nucleic Acids Res.*, **30**, e26.
11. VanBuur, C., Cuypers, H., Lelie, P., Chudy, M., Nübling, M., Melsert, R., Nabbe, A. and Oudshoorn, P. (1998) The NucliSens™ Extractor for automated nucleic acid isolation. *Infusionsther. Transfusionsmed.*, **25**, 147–151.
12. Witt, D.J., Kemper, M., Stead, A., Ginocchio, C.C. and Caliendo, A.M. (2000) Relationship of incremental specimen volumes and enhanced detection of human immunodeficiency virus type 1 RNA with nucleic acid amplification technology. *J. Clin. Microbiol.*, **38**, 85–89.
13. Dudewicz, E.J. and Mishra, S.N. (1988) *Modern Mathematical Statistics*. John Wiley & Sons, New York, p. 264.

## APPENDIX

### Truncated Poisson probabilities

Let  $X$  be a random variable obtained from a Poisson distribution with parameter  $\lambda$  under the restriction that its realization is positive. In this case, the probability distribution function is  $\Pr(X = x) = (e^{-\lambda}\lambda^x) / [(1 - e^{-\lambda})x!]$ . Straightforward calculus now shows that the expectation  $\mu_x = \lambda / (1 - e^{-\lambda})$  and the variance  $\sigma_x^2 = \lambda[1 - (\lambda + 1)e^{-\lambda}] / (1 - e^{-\lambda})^2$ . General statistical theory reveals that approximations for the expectation and variance of some function  $g(\cdot)$  of the random variable  $X$  can be computed (see for example 13):

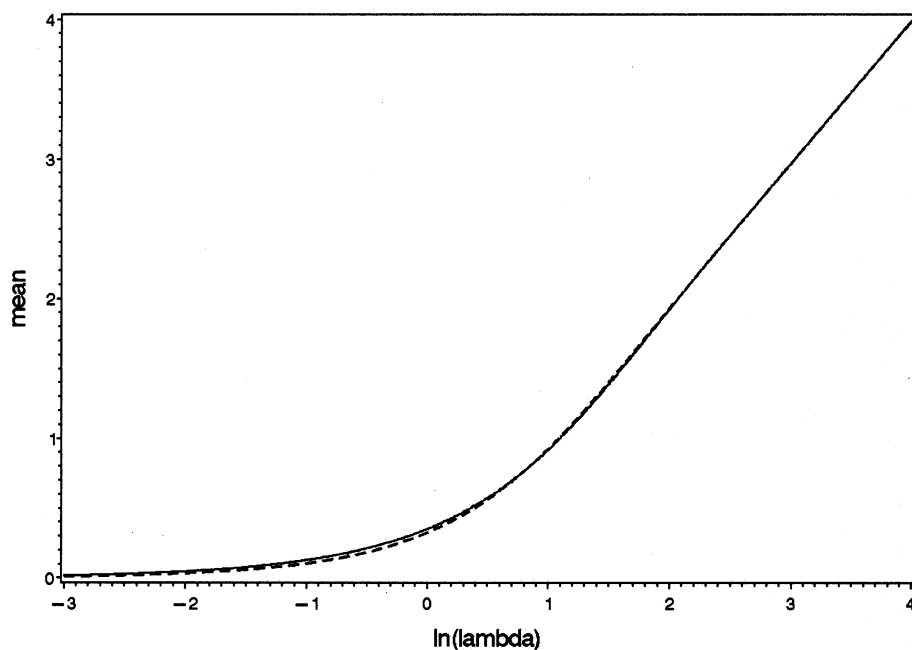
$$\begin{cases} E[g(X)] \approx g(\mu_x) + (1/2)\sigma_x^2 g''(\mu_x) \\ S_{[g(X)]^2} \approx [g'(\mu_x)]^2 \sigma_x^2 \end{cases}$$

implying for the logarithmic transformation  $g(x) = \ln(x)$ :

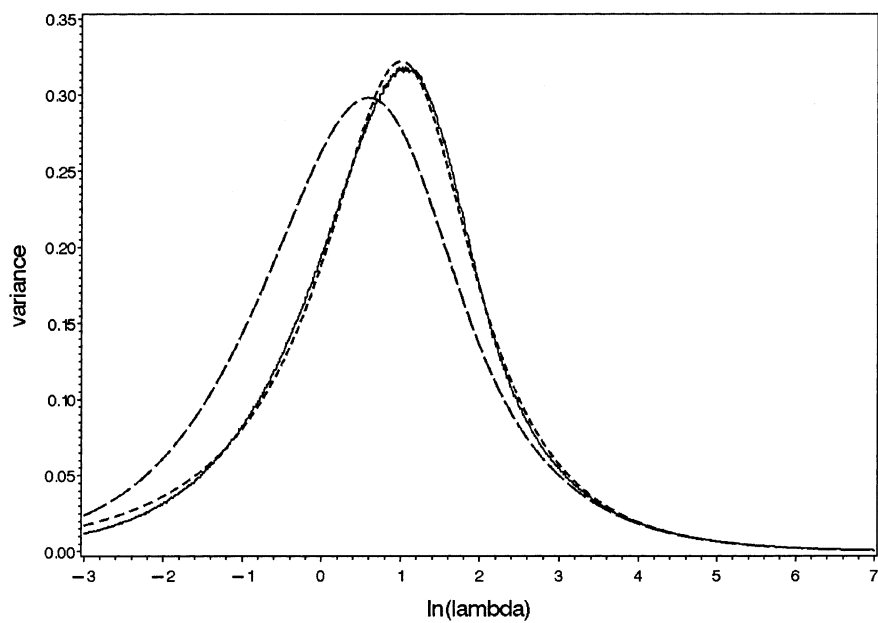
$$\begin{cases} E[\ln(X)] = \ln[\lambda/(1 - e^{-\lambda})] + [(\lambda + 1)e^{-\lambda} - 1]/2\lambda \\ S_{[\ln(X)]^2} = [1 - (\lambda + 1)e^{-\lambda}]/\lambda \end{cases} \quad \mathbf{A1}$$

Large-scale computer simulations were performed to study the validity of these equations. To this end, values were chosen for the Poisson parameter  $\lambda$  by taking values for  $\ln(\lambda)$  between  $-3$  and  $+7$  with increments of  $0.01$ . For each  $\lambda$ , a total of  $10^5$  observations were drawn from truncated Poisson distributions, and the mean and variance of the logs of these random values were computed. The summary statistics so obtained as well as the approximations according to equation **A1** are presented in Figures 4 and 5. It is clear that the approximation for the mean is fairly accurate, but the approximation for the variance fails. Therefore, an alternative equation was searched for. There was no need to define a firm theoretical basis; any equation that is manageable and yields a reasonably good approximation would suffice. Empirical study of the observed curve suggests to use the equation as given in equation **A2**, in which the actual values of the parameters were optimized by using non-linear regression procedures. The quality of this approximation is also presented in Figure 5. This equation is used in the current paper.

$$S_{[\ln(X)]^2} = \frac{1}{\lambda - \vartheta_1 + (\vartheta_2/\lambda^{\vartheta_3})} \quad \text{with} \quad \begin{cases} \vartheta_1 = 3.5765 \\ \vartheta_2 = 7.8346 \\ \vartheta_3 = 0.68208 \end{cases} \quad \mathbf{A2}$$



**Figure 4.** Approximation of the mean of the logarithm of a truncated Poisson variable with equation A1 (dashed line). The solid line reflects the simulation results.



**Figure 5.** Approximation of the variance of the logarithm of a truncated Poisson variable with equations A1 (long dashes) and A2 (short dashes). The solid line reflects the simulation results.