The cost-effectiveness of expanded HIV screening in the United States

Aaron Lucas and Benjamin Armbruster

Objective: The current Centers of Disease Control and Prevention (CDC) guidelines from 2006 recommend a one-time test for low-risk individuals and annual testing for those at high risk. These guidelines may not be aggressive enough, even for those at low risk of infection, due to the earlier initiation of HAART and a movement towards a test-and-treat environment. We evaluated the optimal testing frequencies for various risk groups in comparison to the CDC recommendations.

Methods: We build a deterministic mathematical model optimizing the tradeoff between the societal cost of testing and the benefits over a patient's lifetime of earlier diagnosis.

Results: Under a test-and-treat scenario with immediate initiation of HAART, the optimal testing frequency is every 2.4 years for low-risk (0.01% annual incidence) individuals; every 9 months for moderate risk (0.1% incidence) individuals; and every 3 months for high-risk (1.0% incidence) individuals. The incremental cost-effectiveness of the optimal policy is $36,342/quality-adjusted life-years (QALY) for low-risk individuals and $45,074/QALY for high-risk individuals compared with 20-year and annual testing, respectively.

Conclusion: The current CDC guidelines for HIV testing are too conservative, and more frequent testing is cost-effective for all risk groups.

Introduction
Frequent HIV testing in healthcare settings is an effective method for identifying new HIV infections before substantial declines in CD4\(^+\) T-lymphocyte counts occur [1–4]. Avoiding late diagnosis allows for an earlier initiation of HAART, which substantially reduces HIV transmission in addition to increasing survival [5,6]. There is a growing consensus that initiating HAART immediately after detection is beneficial regardless of a patient’s CD4\(^+\) T-lymphocyte count. In fact, according to the recent US Department of Health and Human Services (DHHS) guidelines, HIV-positive individuals are eligible for antiretrovirals, no matter their CD4\(^+\) T-lymphocyte count [7]. Early detection of HIV through frequent testing also promotes risk reduction in previously undiagnosed HIV-positive patients by decreasing partnership acquisition rates and increasing safe sex practices with perceived HIV-negative partners [8].

In 2006, the Centers for Disease Control and Prevention (CDC) released recommendations on universal voluntary opt-out HIV testing in healthcare settings for adolescents and adults [9]. The CDC recommended that all adolescents and adults who have not been tested before should be offered a test and that high-risk individuals should be tested at least annually. The recommendations consider high-risk individuals to be IDUs, and their sex partners, sex workers, sex partners of HIV-infected persons and individuals who themselves or their sex partner have multiple partners. In addition, the CDC recommended in 2011 that it would likely be beneficial for MSM to screen every 3–6 months [10].
The 2006 CDC recommendations are based on articles by Walensky et al. [11], Sanders et al. [3] and Paltiel et al. [2] examining the cost-effectiveness of routine HIV testing in the out-patient setting. However, since the publication of these articles in 2005, the guidelines for initiating HAART have changed substantially. The International AIDS Society–USA Panel recommends initiating HAART at a CD4⁺ T-lymphocyte count of 500 cells/µL [12]. Given the accelerated rate at which new, safer and more effective drugs are entering the market and the recent finding that HAART reduces the risk of transmitting the virus by 96%, it is not unreasonable to conceive of recommendations to initiate treatment immediately after diagnosis in the near future [6]. In addition, the study by Paltiel et al. [2] assumed all individuals background test every 5 years. They state that this agrees with the observation in the medical community that it takes on average 5 years from infection to detection. However, testing every 5 years with a constant probability of background testing implies approximately 2.5 years from infection to detection.

Our goal is to determine whether HIV testing in healthcare settings should be more aggressive than currently recommended. Although we assume a more liberal treatment scenario in which a patient enters treatment immediately after detection, we consider in the sensitivity analysis waiting to initiate treatment until the CD4⁺ T-lymphocyte count drops to 500 cells/µL. To investigate this, we use a simple mathematical model to evaluate the optimal frequency of HIV testing under various incidence rates, which serve as proxies for risk groups. Our analysis includes the monetary costs and health costs from both the individual tested and from preventing secondary infections.

Materials and methods

Our analysis uses quality-adjusted life-years (QALYs) as the measure of effectiveness. We seek to find an optimal time between HIV tests, which minimizes the total lifetime discounted societal monetary and health costs incurred from an individual using a cost–benefit analysis. For that purpose, we convert QALYs into health costs by using $, the dollar value for one QALY. We later use actual QALYs to calculate cost-effectiveness ratios.

Table 1 lists the main parameters of the model [13–15]. We let $q$ denote the annual discount rate for monetary and health costs; $τ$ denote the duration a person is sexually active during their lifetime; $g$ denote the cost-effectiveness threshold; and $C$ denote the cost of an HIV test. Monetary costs are inflated to 2010 dollars using the medical care component of the CPI-U [16]. For our cost–benefit analysis, we use a baseline value of $q$ equal to $163,880 equal to one QALY [13]. As the value $q$ equal to $50,000 is used abundantly in the medical literature, we consider this value in the sensitivity analysis [17]. The baseline cost of a rapid HIV antibody test is $C$ equal to $16 [14], and in the sensitivity analysis, we consider a cost of $56 [18] and $42 [18]. We let $h$ denote the incidence of HIV infection for the population of interest. We consider annual rates of 0.01, 0.1 and 1% to serve as proxies for low, moderate and high-risk groups, respectively. Recent estimates for HIV incidence in the United States among those aged 13 years and older peg the overall rate at 0.02% [19,20], with the incidence rate among whites at 0.01% and among blacks at 0.07% [20]. The incidence among MSM is even higher. MSM compose the largest group of new HIV infections [19] and have an average incidence of 0.7% [21]. In some urban populations, MSM incidence even reaches 2–3% [22].

We now define the total discounted net lifetime monetary, $F_1(t)$, and health costs in QALYs, $F_0(t)$, of detecting HIV $t$ years after infection in an individual. These functions capture the costs not only to the person diagnosed but also those from secondary infections. We are particularly interested in the slopes of these functions, $f$ and $h$, respectively, because they describe the monetary and health costs of delaying diagnosis by 1 year (and thus delaying treatment in a test-and-treat environment). We combine these into a total marginal cost of detection (MCD) using the dollar amount of one QALY, $q$, so that the MCD = $f$ = $f$ + $gh$. Our assumption that $f$ and $h$ are independent of $t$ [i.e. that $F_1(t)$ and $F_0(t)$ are close to linear] is reasonable due to the lack of better data and because we expect the rate of secondary infections to be constant during the long, chronic stage of untreated HIV infection. Although the acute, primary stage of HIV infection accounts for a disparate large amount of secondary infections, we assume that standard antibody tests rolled out during a test-and-treat policy likely do not detect HIV infection before the chronic stage begins. Using Supplementary Table 1, http://links.lww.com/QAD/A277, we estimate baseline values of these key parameters, $f = -5304$ and $f = 0.41$ QALYs, indicating that delaying the detection of an individual HIV infection

<table>
<thead>
<tr>
<th>Table 1. Baseline parameter values*.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
</tr>
<tr>
<td>Discount rate, $r$ (%) [15]</td>
</tr>
<tr>
<td>Duration of sexually active life, $τ$ (years)</td>
</tr>
<tr>
<td>Test cost, $C$ [14]</td>
</tr>
<tr>
<td>Annual incidence rate for low-risk group, $h$ (%)</td>
</tr>
<tr>
<td>Annual incidence rate for medium risk group, $h$ (%)</td>
</tr>
<tr>
<td>Annual incidence rate for high-risk group, $h$ (%)</td>
</tr>
<tr>
<td>$$/QALY, g$ [13]</td>
</tr>
<tr>
<td>Incremental cost of detecting HIV 1 year after infection</td>
</tr>
<tr>
<td>Monetary cost, $f_c$ (QALY)</td>
</tr>
<tr>
<td>Health cost, $f_h$ (QALY)</td>
</tr>
</tbody>
</table>

QALY, quality-adjusted life-year.

*All costs are in 2010 dollars.
by 1 year saves society $5304 and costs it 0.41 QALYs, yielding a total MCD of $61 293 for each year of delay.

We assume that individuals are initially HIV-negative, a conservative assumption that underestimates the value of testing; that the annual probability of infection, \( h \), remains constant; that the testing is close to 100% accurate [23–25]; and that HAART reduces secondary transmission to a negligible rate [6]. Another key assumption is that HIV-positive individuals may begin HAART immediately after diagnosis. In the sensitivity analysis, we examine a scenario in which HIV-positive individuals only begin HAART at a CD4 T-lymphocyte count of 500 cells/\( \mu l \). In that scenario, we assume that between the time of diagnosis and the initiation of HAART, the individual receives no direct benefit from knowing their status but reduces their risky behaviour by 20% [26,27]. We use a CD4 T-lymphocyte count of 750 cells/\( \mu l \) immediately after seroconversion [28,29], with an average yearly CD4 cell count decline of 75 cells/\( \mu l \) [28,30–32] to estimate that the time from infection to a CD4 T-lymphocyte count of 500 cells/\( \mu l \) is 3.3 years. As we seek to capture all the relevant costs, we do not include additional background testing in our model unlike the one by Paltiel et al. [2].

A concern with implementing an effective test-and-treat scenario is timely linkage to care and sustained engagement in care for HIV-diagnosed individuals [33]. Sustained engagement includes not just linking to HIV care, but full retention in care and adequate adherence to antiretroviral medication. A recent meta-analysis of linkage to care studies conducted after 2003 found that only 72% of those diagnosed as HIV-positive were properly linked to care and only 42% had multiple HIV medical care visits during a specified assessment interval [34]. Moreover, even after the facilitation of a connection between an HIV patient and provider, the patient’s long-term retention into care and the patient’s adherence to antiretroviral treatment remains uncertain. To model this uncertainty, we create a parameter \( p \) for the probability with which a patient is linked to and fully engaged in HIV care soon after diagnosis. This probability is hard to pinpoint, and so, we examine \( p \) at a low value of 0.3 and a higher value of 0.9 in the sensitivity analysis; we fix \( P \) equal to 1 at baseline. The required modification of the model is detailed in the supplement.

We let \( T \) denote the time between HIV tests. We examine the trade-off between the cost of testing for HIV and the cost of delaying detection. Our goal is to seek the time between HIV tests, \( T^* \), which minimizes \( g(T) \), the expected societal discounted lifetime combined monetary and health cost per person. The details of calculating \( T^* \) are provided in the supplement. Nevertheless, there is also a simple formula (see the supplement for details, http://links.lww.com/QAD/A277) to approximate \( T^* \) within 3% of its true value:

\[
T^* \approx \sqrt{2/C/(f_h^* h f_t T_1 T_2)} - f_c/f_h.
\]

We may also revert to the original \( f_c \) and \( f_h \) to perform a cost-effectiveness analysis with QALYs as the measure of effectiveness. Using the approximation in Equation (1) (details are provided in the supplement, http://links.lww.com/QAD/A277), we can approximate the incremental cost-effectiveness ratio of testing every \( T_2 \) years instead of every \( T_1 \) years as

\[
2C/(h f_h T_1 T_2) - f_c/f_h.
\]

We use a 20-year testing frequency as our reference group in our analysis to serve as a proxy for testing once in a lifetime, as 20 years is half of our baseline sexual lifespan.

**Results**

Figure 1 shows the optimal time between tests, \( T^* \), as a function of the annual incidence rate, \( h \). In a low-risk population with an annual incidence of 0.01%, the optimal testing frequency is approximately every 2.4 years. In medium and high-risk populations with an annual incidence of 0.1 and 1%, the optimal testing frequency is approximately every 9 and every 3 months, respectively.

Figure 2 shows the incremental cost-effectiveness of different testing frequencies compared with testing every 20 years for low, medium and high-risk populations. We use a 20-year testing frequency as our reference group to serve as a conservative proxy for a once in a lifetime test. At $100 000 per QALY, annual testing is cost-effective for medium and high-risk groups, whereas every 3 years is cost-effective for low-risk groups. At $50 000 per QALY,
testing low-risk groups every 5 years is cost-effective. For the high-risk group, the ICER is almost constant until we reach an annual testing frequency, whereas for the low-risk group, the ICER increases rapidly as the testing frequency increases. Furthermore, moving from testing every 20 years (a conservative proxy for the current policy for low-risk individuals) to testing with our model’s optimal testing frequency of every 2.4 years results in a gain of 0.006 QALYs at a cost of $213.68 per person. Similarly, testing high-risk individuals every 3 months compared with annual testing yields a gain of 0.03 QALYs at a cost of $1357 per person.

The approximation in Equation (1) provides a simple description of how the optimal testing frequency changes under different scenarios. For example, doubling the cost of an HIV test causes the optimal time between tests to increase by a factor of approximately $\sqrt{2} \approx 1.4$, whereas doubling the MCD, $f$, for the same incidence rate reduces the optimal time between tests by a factor of $\frac{1}{\sqrt{2}} \approx 0.7$. Table 2 shows the optimal testing frequency under different scenarios. The optimal testing frequency is relatively unchanged if we use vastly different sexual lifespans; if we increase the average annual cost of managing HIV infection from $23,007 per year to $25,000; or if we decrease the loss in life-years and QALYs from delaying treatment initiation 1 year from 0.35 to 0.2. The optimal time between tests drops by one-third if we decrease the HIV-positive life expectancy from 37 to 18 years. In addition, the optimal time between tests increases by roughly two-thirds if the cost of testing increases to $42 from $16 and by a half if the rate secondary cases falls to 0.02 per infected person per year from 0.04. The optimal time between tests approximately doubles if we decrease the cost-effectiveness threshold to $50,000 per QALY and almost doubles if we delay the
Table 2. Optimal time between tests, $T^*$, in years by risk group$^a$.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>0. Baseline, $f_c = $61 293</td>
<td>2.4</td>
<td>0.7</td>
<td>0.2</td>
</tr>
<tr>
<td>1. Decreasing the duration of sexually active life, from 40 to 20 years</td>
<td>2.5</td>
<td>0.7</td>
<td>0.2</td>
</tr>
<tr>
<td>2. Increasing the duration of sexually active life, from 40 to 60 years</td>
<td>2.4</td>
<td>0.7</td>
<td>0.2</td>
</tr>
<tr>
<td>3. Increasing the annual cost of HIV care from $23 007 to $25 000 ($f_c = -$5764, $f_h = 0.41)</td>
<td>2.4</td>
<td>0.7</td>
<td>0.2</td>
</tr>
<tr>
<td>4. Decreasing the life-years and QALYs lost by delaying HAART by 1 year from 0.33 to 0.2 ($f_c = -$4148, $f_h = 0.36)</td>
<td>2.5</td>
<td>0.8</td>
<td>0.2</td>
</tr>
<tr>
<td>5. Decreasing HIV-positive life expectancy from 37 to 18 years ($f_c = -$15 080, $f_h = 0.76)</td>
<td>1.7</td>
<td>0.5</td>
<td>0.2</td>
</tr>
<tr>
<td>6. Increasing the test cost, $C = $42</td>
<td>4.0</td>
<td>1.2</td>
<td>0.4</td>
</tr>
<tr>
<td>7. Increasing the test cost, $C = $56</td>
<td>4.4</td>
<td>1.4</td>
<td>0.4</td>
</tr>
<tr>
<td>8. Decreasing the rate of secondary cases from 4 to 2% a year ($f_c = -$15 504, $f_h = 0.26)</td>
<td>3.6</td>
<td>1.1</td>
<td>0.3</td>
</tr>
<tr>
<td>9. Initiating treatment at CD4 cell count of 500 cells/μL instead of immediately</td>
<td>5.0</td>
<td>1.6</td>
<td>0.5</td>
</tr>
<tr>
<td>10. Only 30% successful linkage and retention in care</td>
<td>4.4</td>
<td>1.3</td>
<td>0.4</td>
</tr>
<tr>
<td>11. Only 90% successful linkage and retention in care</td>
<td>2.5</td>
<td>0.8</td>
<td>0.2</td>
</tr>
<tr>
<td>12. Decreasing the cost-effectiveness threshold, $q = $50 000/QALY</td>
<td>5.0</td>
<td>1.5</td>
<td>0.5</td>
</tr>
<tr>
<td>13. Marginal cost of detecting (MCD) infection 1 year later, $t = $30 000</td>
<td>13.3</td>
<td>3.3</td>
<td>1.1</td>
</tr>
</tbody>
</table>

QALY, quality-adjusted life-year.

$^a$All costs are in 2010 dollars. Low, medium and high-risk groups have annual incidences of 0.01, 0.1 and 1%, respectively.

initiation of treatment until the CD4$^+$ T-lymphocytes count falls below 500 cells/μL. With a 30% rate of successful linkage and retention in care, the time between tests becomes 4.4, 1.3 and 0.4 years for those at high, moderate and low-risk, respectively. The closest we get to reproducing the testing frequencies recommended by Paltiel et al. [2] and Sanders et al. [3] is by assuming an MCD of $3000, significantly lower than the baseline value of $61 293.

We lastly compare the aggregate societal gain in moving from the CDC recommendations to our model's optimal testing frequency [approximated by Equation (2)]. The gain in moving from testing every 20 years (a conservative proxy for the current the CDC's recommendations) to the optimal testing frequency of every 2.4 years for one million low-risk individuals in the United States would be 5880 QALYs at a cost of $213.7 million or $36 342 per each additional QALY. The CDC estimates that MSM account for 2% of the United States population, or roughly 6 million people [30]. Assuming a 1.0% incidence for MSM, moving from the 2006 recommendations of annual testing to testing every 3 months would avert a loss of 180 636 QALYs in this population at a cost of $8.1 billion or $45 074 per each additional QALY.

Discussion

We used a simple mathematical model to calculate optimal HIV testing frequencies for different risk groups using HIV incidence rates as proxies. This model focuses on the tradeoff between delaying detection of HIV and the cost of testing. In our baseline calculations, we assume that an HIV-positive patient undergoes treatment immediately after diagnosis. Under this scenario, we find that even low-risk individuals may benefit from testing more often than every 3 years. This finding stands out from the once in a lifetime test that the CDC recommends. We show that moderate risk individuals should test every 9 months. This finding supports the notion that the CDC should address those falling between high and low risk, as these individuals are not directly touched upon in the 2006 recommendations. We finally find that the high-risk individuals should test every 3 months. The last recommendation agrees with the recent update from the CDC advising MSM to consider testing every 3–6 months [10], and our findings suggest a large societal health benefit from incorporating this update into the official CDC HIV testing recommendations. Overall, our model explicitly demonstrates that a more aggressive shift towards more frequent testing would save many QALYs at a relatively low cost to society.

Our one-way sensitivity analysis shows our findings to be relatively robust. Particularly encouraging was the fact that the remaining sexual lifespan did not significantly change our results, implying that age is not an important factor. The current recommendations of the International AIDS Society - USA Panel are to initiate treatment when the CD4$^+$ T-lymphocytes count falls below 500 cells/μL, though it is conceivable that that in the future, the recommendation will be to initiate treatment immediately after diagnosis. If we assume that treatment may only be initiated at a CD4$^+$ T-lymphocytes count of 500 cells/μL (approximately 3.3 years after infection), we still find an optimal testing frequency of every 5 years, which is much more frequent than than the once in a lifetime test currently recommended for low-risk individuals. Even in a setting with a low rate of linkage and retention in care, we still find an optimal testing interval of every 4.4 years for those at low risk and 1.3 years for those at moderate risk. Thus, even in settings with poor rates of follow-up, those at moderate or high risk should be tested close to annually.
Our model finds it optimal to test much more frequently than what Paltiel et al. [2] and Sanders et al. [3] suggest. In addition, our ICERs for testing every 5 years instead of every 20 years are significantly lower than the ICERs in the studies by Paltiel et al. [2] and Sanders et al. [3] for testing every 5 years instead of once (they did not consider every 20 years). This implies a much lower MCD of $3000, compared with our baseline value of $61,293. Unfortunately, due to the complexities of the models in Paltiel et al. [2] and Sanders et al. [3], we cannot pinpoint a reason for the discrepancy.

Now, we turn to the limitations of our model. The first limitation is that similar to the earlier models, we assume that each individual experiences a constant hazard rate of becoming infected. However, the risk of infection in a monogamous partnership may decrease over time, as it becomes more and more certain that both individuals are likely to be HIV negative. Thus, the assumption of a constant hazard rate may be invalid when the partner turnover is very low. Thus, testing recommendations for very low-risk groups may be better phrased not as a test every T years but as a test every n partners. The assumption of a constant hazard rate also becomes invalid when a person’s risk profile changes over their sexually active lifetime (e.g. when they enter a long-term monogamous relationship). However, one can always change the person’s recommended testing frequency as their risk profile changes. The second limitation is that our calculation of the incremental net monetary and health costs did not capture any symptomatic case finding. However, symptoms are unlikely to develop before detection by routine testing with the frequencies we suggest. Third, we do not address costs associated with a positive HIV test, including initial posttest counselling. These, however, are fixed costs for every newly diagnosed HIV-positive person and thus do not affect the optimal value in our equations. Fourth, we briefly mention the ‘window period’ of standard antibody HIV tests in the introduction that they often do not detect infections in the acute stage (infections within 1–3 months for standard tests). However, this is not a limitation of our model, as adding the window period into our model would only shift our equations by a constant and not affect the MCD and thus would not affect the optimal testing frequencies. Fifth, we acknowledge that although our baseline assumes a test-and-treat scenario, such a testing environment is not yet a reality in every healthcare setting. However, this scenario is quickly approaching as evidenced by the recent DHHS guidelines of initiating ART regardless of a patient’s CD4+ T-lymphocytes count [7]. Thus, our findings are directly applicable if a particular healthcare setting follows the DHHS guidelines. Finally, although we do take into account the present value of any future costs associated with HIV diagnosis, we acknowledge both the complexity and uncertainty surrounding the long-term projections of these costs in treating HIV-positive individuals. However, it is likely that HIV antibody test costs will decrease as time progresses. Moreover, even if costs associated with HAART increase, the benefits of early diagnosis with respect to extending survival and secondary infections will still persist.

We have discussed the limitations of two studies [2,3] that influenced the 2006 CDC guidelines on routine HIV testing in healthcare settings [9]. Those reasons and the recommendations for earlier initiation of treatment [12] led us to develop a straightforward mathematical model to suggest new policies. Our model uses a multipronged approach. We implemented a cost–benefit analysis to determine the optimal time between HIV tests for individuals by risk group and a cost-effectiveness analysis to demonstrate the improvement in moving from the CDC recommendations to our model’s recommendations. Our results conclusively showed that the 2006 CDC HIV testing recommendations are too conservative, especially for low-risk groups who would benefit from more frequent testing. These results should encourage policymakers and medical professionals to reconsider how often adults and adolescents should be tested for HIV.

Acknowledgements

The authors received no funding for this report. All authors contributed equally to the report. The authors thank Sabina Alistar and Rebecca Wurtz for helpful comments and suggestions.

There was no funding for this report.

The study was presented at the 33rd Annual Meeting of the Society for Medical Decision Making on 24 October 2011 in Chicago, Illinois, USA.

Conflicts of interest

The authors have no commercial or other association that would pose a conflict of interest.

References


