

Public Health Consequences of Screening Patients for Adherence to Highly Active Antiretroviral Therapy

*Eric Tchetgen, †Edward H. Kaplan, and ‡Gerald H. Friedland

**Department of Family Medicine and Community Health, Tufts University School of Medicine, Boston, Massachusetts, U.S.A.;*

†*Yale School of Management and Center for Interdisciplinary Research on AIDS, Yale University, New Haven, Connecticut, U.S.A.; and* ‡*AIDS Program, Yale School of Medicine, New Haven, Connecticut, U.S.A.*

Summary: Improvements in HIV antiretroviral therapy (ART) have been accompanied by increasing recognition of the importance of adherence to treatment regimens for maximizing patient benefits while minimizing the emergence of drug-resistant virus. Whether clinicians should screen patients for adherence and only administer therapy to those believed likely to adhere has not been resolved. We first examine the implications of data drawn from a recent study reporting physicians' ability to predict whether patients will adhere to highly active antiretroviral therapy (HAART) or not. We then extend previously developed mathematical models of ART to include screening for adherence and focus on resulting drug resistance as well as on HIV and AIDS incidence at the population level. We show that although screening for adherence is likely to reduce the level of drug resistance compared with a policy of treating all HIV patients with HAART, rates of new HIV infections and AIDS cases in the population would likely increase unless screening accuracy is extremely (perhaps implausibly) high. **Key Words:** Adherence—HAART—Drug resistance—AIDS incidence—Mathematical model.

HIV antiretroviral drug therapy (ART) has improved dramatically over the past 10 years. Notably, recent advances in HIV therapeutics have significantly increased disease-free survival time for many infected patients. Highly active antiretroviral therapy (HAART), a combination of three or more ARTs, has proven successful at suppressing HIV viral replication and encouraging immune response in treated patients (1). Recent studies have reported the emergence of drug-resistant strains of HIV associated with the widespread use of these antiretroviral drugs, however (2,3). Nonadherence to HAART has been identified as a major factor associated with drug resistance and therapeutic failure (4–8). With the growing presence of primary and acquired drug-resistant HIV in the United States and other resource-rich countries

where HAART has become the standard of care, clinicians have become increasingly concerned that nonadherence might lead to levels of drug-resistant HIV with clinical and public health consequences.

Recent articles have examined the effects of ART on the spread of HIV and the development of drug-resistant HIV strains (9,10) as well as the role of adherence in reducing the incidence and prevalence of drug resistance among HIV cases (6,7,9). Although these papers were successful in characterizing the association between access to ART, patients' adherence to therapy, and the emergence of drug resistance, they did not discuss the important role of clinicians as decision makers faced with the responsibility of prescribing HAART to patients seeking treatment for HIV. In an effort to prevent the emergence of drug-resistant HIV, clinicians have been encouraged to screen patients for adherence before prescribing HAART and to offer the most potent therapy only to those patients deemed ready to adhere to the prescribed regimen (11,12). It is argued that screening

Address correspondence and reprint requests to Edward H. Kaplan, Yale School of Management, P.O. Box 208200, New Haven, CT, 06520–8200, U.S.A.; e-mail: edward.kaplan@yale.edu

Manuscript received August 17, 2000; accepted October 20, 2000.

for adherence benefits the individual patient and successfully decreases the incidence and prevalence of drug-resistant HIV, thereby ensuring the therapeutic effectiveness of HAART for future generations of HIV patients. The extreme version of this proposed screening policy has recently been characterized as “to treat or not to treat” current seropositive patients (13), reflecting the tension between clinicians’ medical commitment to treat current HIV patients and the strategic desire to preserve the “treatability” of future generations of HIV patients by maintaining a low level of drug-resistant HIV in the population. A more reasonable and currently recommended policy advocates that HAART be delayed until prospective patients are able and ready to adhere to the treatment regimen; such “treat or wait” decisions would be made on the basis of screening HIV patients for adherence.

What has not been studied, however, are the public health consequences of screening for adherence. Physicians faced with the task of prospectively determining which patients are more or less likely to adhere to HAART have yet to develop systematic procedures for making such judgments and thus inadvertently but inevitably commit screening errors. It is the intent of this article to call attention to the public health consequences of such errors. Providing HAART to nonadherent patients leads to an increase in the likelihood of developing drug resistance and a reduction in clinical benefit from what would be expected for an adherent patient. Withholding HAART, even temporarily, from patients who would adhere to HAART implies that viral replication would continue unchecked in such patients. Also, to the extent that HIV transmission to others depends on viral load, secondary transmission of HIV becomes more likely if HAART is withheld (10,14). Given such screening errors, it is conceivable that screening for adherence might simultaneously deliver individual patient benefits while worsening HIV and AIDS incidence at the population level.

This article presents a model to examine the epidemiologic implications of screening HIV patients for adherence to HAART. We first consider data on the precision of physician screening for adherence reported recently by Paterson et al. (15). These data suggest that for those patients studied, clinicians’ judgments regarding whether patients would adhere to HAART or not were statistically independent of whether patients actually adhered. We then join an epidemic model of the transmission dynamics of HIV with a probabilistic model that accounts for the frequency and accuracy with which physicians screen patients for adherence. Our model of screening is akin to a diagnostic test where clinicians judge patients as adherent or not with characteristic de-

grees of accuracy as implied by test sensitivity and specificity. For simplicity, screening sensitivity and specificity are always equal in this model (although they need not be in general). We refer to this common screening sensitivity/specificity as “screening accuracy” and consider the implications of differential screening accuracy on the spread of resistant virus as well as on the overall incidence of HIV and AIDS in the population. The model suggests that as a function of the true level of adherence in the patient population and screening accuracy, screening for adherence may not always be beneficial. In particular, we show that the screening accuracy required to justify screening for adherence not only depends on the level of adherence in the population of HAART users but also depends crucially on whether one considers the reduction of drug resistance, HIV incidence, or AIDS incidence as the desired policy goal.

SCREENING FOR ADHERENCE: EMPIRICAL OBSERVATIONS

Paterson et al. (15) reported the differential outcomes for HAART patients with different levels of adherence as measured by the Medication Events Monitoring System (MEMS). They also asked physicians to predict whether individual patients would adhere or not before the MEMS results were known, where “adherence” was operationalized as taking at least 80% of the prescribed dose of medication. With respect to physician screening accuracy, Paterson et al. (15) reported the following three statistics:

- Of those patients whom physicians predicted would fail to adhere, 51% in fact did adhere (page 26).
- 41% of all patients were misclassified by physicians (pages 26 and 28).
- 28% of all patients were judged to adhere but in fact did not take at least 80% of the prescribed dose (referred to as “overestimation of adherence” on page 28).

Again, interpreting adherence as meeting an 80% cutoff as measured by MEMS, these statistics imply that:

- 47% of all patients both adhered and were predicted to adhere.
- 28% of all patients did not adhere but were predicted to adhere.
- 13% of all patients adhered but were predicted to not adhere.
- 12% of all patients did not adhere and were predicted to not adhere.

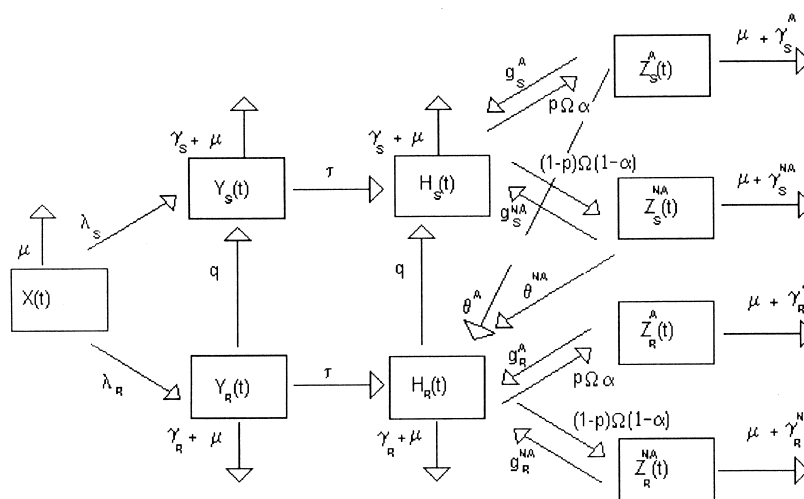
Because 60% (47% + 13%) of all patients adhered, the observed fraction of adherent patients correctly classified by physicians, which can be thought of as the “sensitivity

ity” of screening for adherence, is equal to 78% (47%/60%). The observed fraction of nonadherent patients correctly classified by physicians, which can be thought of as the “specificity” of screening for adherence, was only equal to 30% (12%/40%), however. Thus, although physicians correctly identified 78% of adherent patients, physicians also predicted that 70% of nonadherent patients would have adhered to HAART. Given that these data were derived from 81 patients, there is no statistical difference between the fraction of adherent patients predicted to be adherent (78%) and the fraction of nonadherent patients predicted to be adherent (70%) ($z = 0.88$; $p = .38$). Physicians’ predictions of adherence were thus statistically independent of the actual adherence of patients. Note also that although physicians predicted that about 75% of all patients would adhere (at the 80% MEMS cutoff), the data reveal that only 60% of patients actually did so.

This is a troubling result, for it suggests that using whatever procedures were employed to predict patient

adherence in the study of Paterson et al. (15), physicians’ predictions were noninformative. If it is generally the case that the results of physicians’ judgments and patient adherence status are independent, a policy that supports providing HAART to only those patients who physicians believe will adhere is equivalent to a policy that provides HAART to randomly selected patients. For example, from the study of Paterson et al. (15), where physicians predicted that about 75% of all patients would adhere to medications, a “treat if believed adherent” policy would result in only 75% of patients receiving treatment, even though those denied treatment would be just as likely to adhere as those treated with HAART.

If physicians are not able to discriminate between adherent and nonadherent patients with a reasonable degree of accuracy a priori, there is little reason not to treat all patients so long as the potential benefits of HAART outweigh the risks of developing drug resistance. The greater the overall level of adherence in the patient population, the more likely the overall benefits are to out-



Variables

- $X(t)$: Susceptible
- $Y_s(t)$: Undetected / Drug sensitive
- $Y_r(t)$: Undetected / Drug resistant
- $H_s(t)$: Holding / Drug sensitive
- $H_r(t)$: Holding / Drug resistant
- $Z_s^A(t)$: Adherent / Drug sensitive
- $Z_r^A(t)$: Adherent / Drug Resistant
- $Z_s^{NA}(t)$: Non -adherent / Drug sensitive
- $Z_r^{NA}(t)$: Non -adherent / Drug resistant

Parameters

- Π : Arrival rate
- c : Average number of new sexual partners per year
- μ : Departure rate
- τ : Detection rate
- q : Drug - Resistant to drug - sensitive reversion rate
- γ_s^U : Untreated / drug sensitive AIDS progression rate
- γ_r^U : Untreated / drug resistant AIDS progression rate
- γ_s^A : Adherent / drug sensitive AIDS progression rate
- γ_r^A : Adherent / drug resistant AIDS progression rate
- γ_s^{NA} : Non -adherent / drug sensitive AIDS progression rate
- γ_r^{NA} : Non -adherent / drug resistant AIDS progression rate
- λ_s : Drug sensitive HIV incidence rate
- λ_r : Drug resistant HIV incidence rate
- g_s^A : Adherent / drug sensitive recidivism rate
- g_r^A : Adherent / drug resistant recidivism rate
- g_s^{NA} : Non -adherent / drug sensitive recidivism rate
- g_r^{NA} : Non -adherent / drug resistant recidivism rate
- θ^A : Adherent mutation rate
- θ^{NA} : Non -adherent mutation rate
- Ω : Adherence screening rate
- p : Proportion of adherent patients among newly treated
- α : Adherence screening accuracy

FIG. 1. HIV epidemic model.

weigh the overall risks. In contrast to the implications of Paterson et al.'s data (15), however, we believe that many physicians can differentiate between patients who are or are not likely to adhere to HAART (or could learn to do so with appropriate training). The more relevant questions then become:

- What are the public health consequences of informative approaches to screening patients for adherence to HAART?
- How accurate must physicians' judgments be for a 'treat if believed adherent' policy to improve public health outcomes?

MODEL DESCRIPTION

To help address these questions, we developed an epidemic model representative of a population of gay men as in the recent work of Blower et al. (10). The model flow and parameters are defined in Figure 1, and the equations of the model appear in the appendix. The model divides the population into HIV-susceptible and HIV-infected persons. Susceptible individuals enter the population at a constant rate, while new HIV infections occur with incidence rates that are proportional to the annual number of new sexual partners and transmission probabilities, which depend on whether the infecting virus is sensitive or resistant to HAART. Infected patients are described in terms of their strain of infection (sensitive or resistant to HAART), their treatment status (infected but undetected and hence untreated; infected and detected but currently untreated; or infected, detected, and receiving HAART), and their "true" adherence dichotomized as adherent or nonadherent (reflecting whether the patient would truly adhere if treated with HAART). Patients who are infected and detected but not receiving treatment are said to be in a "holding state." The holding state can serve either as an interim state between seroconversion and initiation of HAART (allowing for physician evaluation to determine which regimen is appropriate) or as a state in which patients who have left HAART for various reasons (e.g., regimen complexity, drug toxicity, development of resistance) can recover with the assistance of their clinician before initiating another round of HAART. All newly detected HIV patients enter the holding state at a constant HIV screening rate, although patients on HAART return to the holding state at a rate that depends on their level of adherence as well as on their strain of infection (sensitive or resistant). Consistent with the literature, we assume that acquired drug resistance emerges among drug-sensitive patients receiving HAART at rates that depend on their level of adherence (6,16). We assume that

among HAART users, new resistant cases are detected immediately and these patients are returned to the holding state for evaluation, after which a new round of HAART can begin.

In the absence of screening for adherence, all patients in the holding state enter HAART at a constant rate. Because primary drug resistance has been observed to revert to drug sensitivity in the absence of therapeutic pressure (2,17), we allow for the possibility that untreated drug-resistant HIV patients may become drug sensitive. We assume that the average incubation time for an infected individual depends on the dominant strain of HIV in his blood, his treatment status, and if on HAART, his true adherence status. In addition to progression to AIDS, men in our model can exit the population for non-HIV/AIDS-related causes.

As stated in the introduction, clinician behavior is incorporated in our model through a parameter reflecting accuracy when screening for adherence. Treating the evaluation of adherence in a manner akin to standard diagnostic testing, we define screening sensitivity as the probability that a clinician correctly identifies a patient as adherent given that the patient actually adheres to a prescribed HAART regimen and screening specificity as the probability that a clinician correctly identifies a patient as nonadherent given that the patient does not adhere to his prescribed HAART regimen. We recognize that adherence is not an "all or nothing" proposition but

TABLE 1. Parameter estimates

Parameter ^a	Adherence favorable	Adherence weakly favorable	Adherence strongly favorable
Π	2133/year	2133/year	2133/year
c	3.4/year	3.4/year	3.4/year
μ	0.033/year	0.033/year	0.033/year
τ	0.5/year	0.5/year	0.5/year
Ω	2/year	2/year	2/year
q	0.1/year	0.1/year	0.1/year
γ_S^U	0.083/year	0.083/year	0.083/year
γ_R^U	0.058/year	0.058/year	0.058/year
γ_S^A	0.031/year	0.032/year	0.027/year
γ_R^A	0.038/year	0.039/year	0.033/year
γ_S^{NA}	0.046/year	0.045/year	0.053/year
γ_R^{NA}	0.057/year	0.055/year	0.065/year
β_S^U	0.1	0.1	0.1
β_R^U	0.05	0.05	0.05
β_S^A	0.026	0.026	0.026
β_R^A	0.025	0.025	0.025
k^S	0.5	0.5	0.5
θ^A	0.076/year	0.135/year	0.012/year
θ^{NA}	0.761/year	0.673/year	0.859/year
g_S^A	0.05/year	0.071/year	0.001/year
g_R^A	0.1/year	0.143/year	0.003/year
g_S^{NA}	0.175/year	0.143/year	0.248/year
g_R^{NA}	0.35/year	0.286/year	0.496/year

^a See Figure 1 and the Appendix for parameter definitions.

maintain adherence as a dichotomous variable for clarity in analysis (as other researchers such as Paterson et al. [15] have done). Simplifying matters further, we assume that clinicians are symmetric screeners in that their sensitivity and specificity values are identical and equal to

what we call screening accuracy. In the model, physicians screen patients for adherence when the latter are in the holding state. Those patients who screen as nonadherent remain in the holding state but can be screened again at a later time.

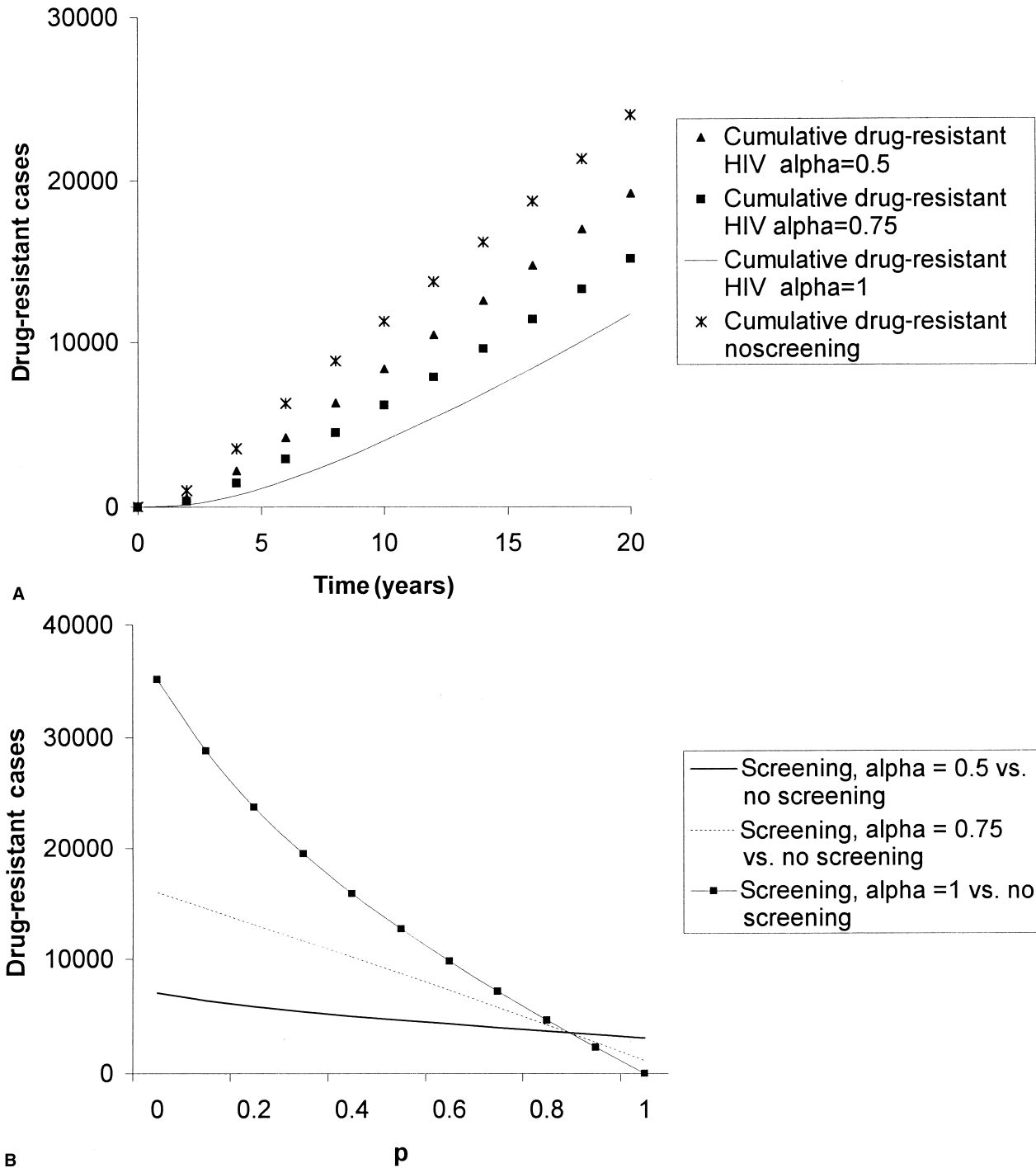


FIG. 2. (A) Base case drug-resistant cases ($p = .5$). (B) Base case averted drug-resistant cases.

EVALUATING THE IMPACT OF SCREENING FOR ADHERENCE

We evaluate the population impact of screening for adherence by contrasting two scenarios, one where physicians screen patients in the holding state for adherence before initiating HAART and treat only those patients deemed to be adherent and the other where all patients receive HAART (and no attempt is made to screen for adherence). Using our model, we estimate the cumulative number of drug-resistant cases prevented, cumulative number of HIV infections prevented, and cumulative number of AIDS cases averted over time on account of screening for adherence. We repeat these comparisons for possible screening accuracies ranging from 50% to 100% and for true population adherence levels ranging from 0% to 100%. We have selected 50% as a lower

limit for screening accuracy, because in an informative screening process, physicians should not perform worse than the results one would obtain from coin tossing. In this manner, we are able to estimate the number of resistant infections, total infections, and AIDS cases averted by screening for adherence as a function of physician screening accuracy and the true fraction of patients in the population who are adherent.

PARAMETERIZING THE MODEL

The San Francisco HIV epidemic has been studied extensively, enabling the estimation of several parameters corresponding to the epidemic in the gay community there. Blower et al. (10) accounted for the uncertainty in the remaining unknown parameters via a Latin hypercube sampling uncertainty analysis, which is a type

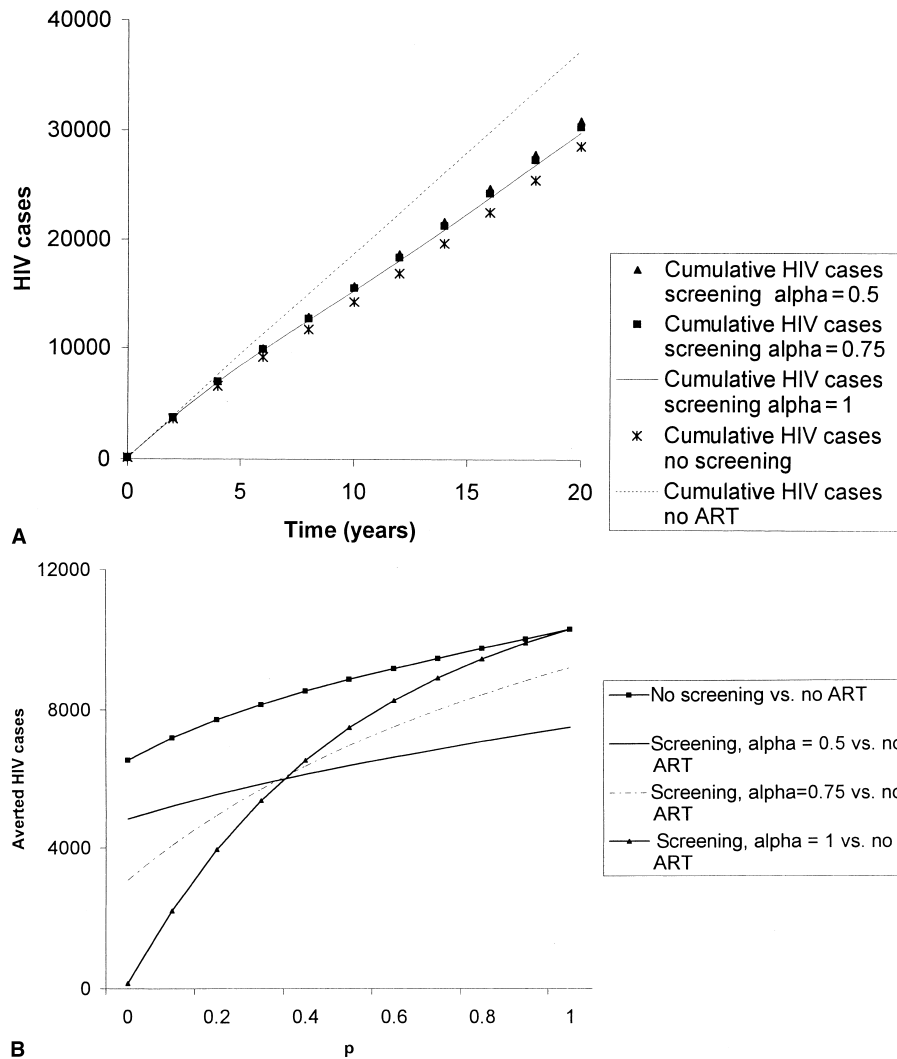
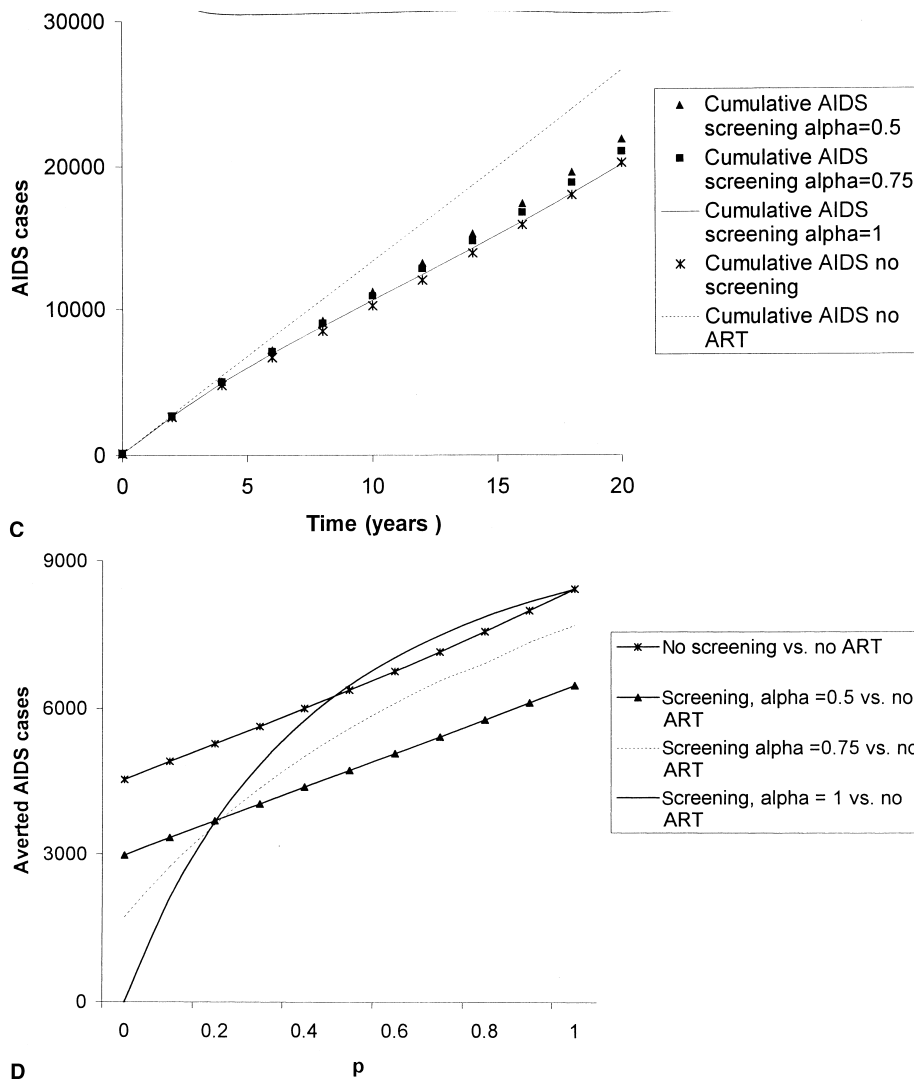


FIG. 3. (A) Base case cumulative HIV cases ($p = .5$). (B) Base case averted HIV cases. (C) Base case cumulative AIDS cases ($p = .5$). (D) Base case averted AIDS cases.

of stratified Monte Carlo sampling. To establish a range of predictions for the San Francisco epidemic, these investigators assigned probability distributions over unknown parameters. Based on these probability distributions, we derived average parameter estimates for use in our study. The parameter values employed appear in Table 1.

The effect of adherence on the effectiveness of HAART is reflected in differential AIDS progression rates, viral mutation rates, and rates of withdrawal from treatment for adherent and nonadherent patients. Current clinical research suggests that the risk of developing drug resistance, progressing to AIDS, or withdrawing from therapy is higher for nonadherent individuals. Our parameter choices reflect these beliefs. Similar to Zaric et al. (9), we assume that nonadherent patients progress to

AIDS at a rate that is 50% higher than that of adherent patients, develop drug resistance at 10 times the rate of adherent patients, and withdraw from therapy at a rate that is 3.5 times higher than their adherent counterparts. Recent epidemiologic studies have reported that approximately 60% of current HAART users adhere to their drug regimen (18–21). As explained in the appendix, joining this observation with our assumed adherence differentials and the mean parameter values from Blower et al.'s study (10) enables estimates of the mean incubation time, emergence rate of resistant virus, and rate of withdrawal from therapy for treated adherent/drug-sensitive cases, nonadherent/drug-sensitive cases, adherent/drug-resistant cases, and nonadherent/drug-resistant cases, respectively. The numeric values employed are summarized in the Appendix.



RESULTS

Figure 2A shows the time trajectory of cumulative drug-resistant cases for a population where 50% of newly treated patients truly adhere to HAART. The figure reports model-derived results in the absence of screening for adherence as well as when screening with accuracies of 50%, 75%, and 100% is employed. Screening for adherence effectively reduces the rate at which cumulative drug-resistant cases emerge over a 20-year time horizon. Figure 2B suggests that the total number of drug-resistant cases averted over 20 years via screening for adherence declines with the true level of adherence in the population.

Again, assuming that 50% of newly treated patients are truly adherent, our model suggests that cumulative HIV and AIDS cases increase over time at a slower rate when HAART is employed than without HAART, irrespective of the accuracy of screening for adherence (Figs. 3A, 3C). When clinicians screen for adherence, however, screening inaccuracy significantly reduces the preventive effectiveness of HAART. In fact, Figures 3A and 3C show that even with perfect screening accuracy, cumulative new HIV infections always remain greater than without screening. Although screening with perfect accuracy reduces total AIDS cases averted over 20 years, the cumulative number of AIDS cases in the absence of screening remains lower for most of the 20 years.

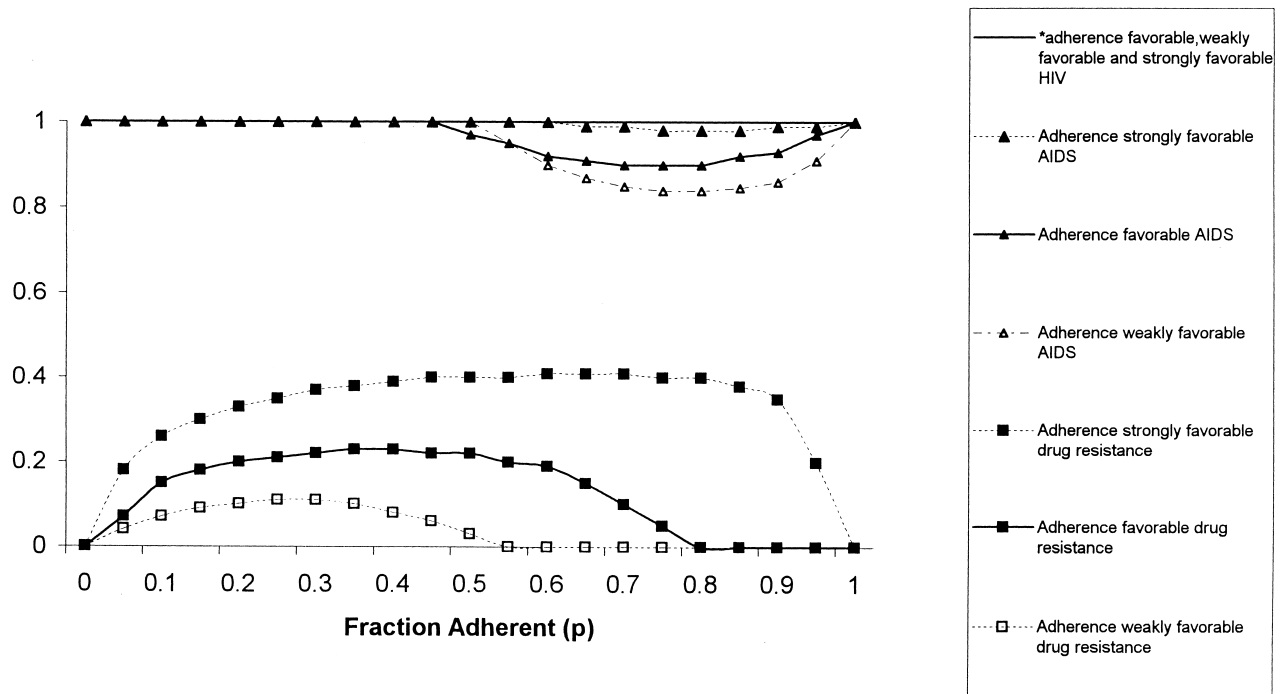
Moreover, as clinicians' screening accuracy drops from 100% to 50%, the growth in cumulative HIV and AIDS cases increases significantly. Figures 3B and 3D reinforce this result by showing the total number of HIV and AIDS cases averted by either screening or not screening before prescribing HAART relative to not offering HAART at all over 20 years as a function of the level of adherence among newly treated patients and clinicians' screening accuracy. Simply treating all patients with HAART without screening for adherence always averts more HIV and AIDS cases than screening with low or medium accuracy. Assuming that clinicians are perfect screeners, it also seems that screening for adherence effectively prevents AIDS cases only if the true level of adherence in the population exceeds 50% (see Fig. 3D). In that same figure, note that screening becomes less effective at preventing AIDS cases as the proportion of patients truly adherent declines.

These results suggest that screening for adherence is not always beneficial in terms of averting drug-resistant cases, new HIV infections, or AIDS cases. To show this, we have estimated the threshold screening accuracy required to reduce the number of HIV, AIDS, or drug-resistant HIV cases over 20 years when screening for

adherence is employed (in comparison to simply treating everyone) for a given level of adherence in the patient population. Figure 4 reports these threshold accuracies as functions of true adherence levels in the patient population. These curves elegantly summarize feasibility regions for adherence screening with respect to each of the three epidemiologic measures we have considered. For whichever of our three measures is of interest, at a given true level of adherence in the patient population, screening for adherence is only justified if the screening accuracy that could be achieved by physicians exceeds the value reported by the threshold curve for the corresponding epidemiologic measure. For instance, if 50% of newly treated patients would truly adhere to HAART, a screening accuracy of only 20% would serve to avert drug-resistant cases under the assumptions we have made thus far (see "Adherence favorable drug resistance" curve in Fig. 4). Coin tossing would be more than sufficient to achieve this level of accuracy. To avert AIDS cases under the same assumptions, however, screening for adherence would have to be at least 90% accurate (see "Adherence favorable AIDS" curve in Fig. 4).

To further investigate the notion of threshold accuracies, we created two additional scenarios termed "Adherence strongly favorable" (which magnifies the benefits of adherence over nonadherence) and "Adherence weakly favorable" (which diminishes the benefits of adherence over nonadherence). Although the precise positioning of the threshold curves shifts, the qualitative properties of these curves are similar. It is clear that the accuracy thresholds are different depending on whether one is interested in preventing drug-resistant cases, AIDS cases, or HIV cases. Preventing drug-resistant infections over 20 years can be accomplished with rather inaccurate screening (accuracies of 50% or less, which again could be achieved via coin tossing), although preventing AIDS cases requires much greater accuracy over the same time period (accuracies of at least 80%). Moreover, for all the scenarios we considered, it seems that HAART with adherence screening is never as effective as simply always administering HAART in preventing HIV infections, even if screening is perfectly accurate.

To further assess the preventive effectiveness of perfect screening for adherence in our "Adherence favorable" scenario, we estimated the time required for the cumulative number of new HIV infections, AIDS cases, or drug-resistant infections to drop below the cumulative number of cases that would have occurred had all patients received HAART (i.e., in the absence of screening) as a function of the true level of adherence in the population. The results are shown in Figure 5. Although per-



*Scenarios where treating everyone is preferred to perfect screening are set to $\alpha=1$

FIG. 4. Threshold screening accuracies.

fect screening presents immediate benefits in the prevention of drug resistance, the time until new HIV infections or AIDS cases are actually prevented increases drastically as the true level of adherence in the population declines. For example, if only 25% of newly treated patients adhere to HAART, it could take up to 25 years before the cumulative number of AIDS cases associated with perfect screening is lower than the cumulative number of AIDS cases in the absence of screening. Moreover, it seems that HIV prevention benefits might never materialize, in that the model suggests that a century would pass before the cumulative number of new HIV infections with perfect screening is surpassed by the cumulative number of new HIV infections in the absence of screening for adherence.

DISCUSSION

Our simulations suggest that the epidemiologic benefits of screening HIV patients for adherence to HAART vary dramatically depending on whether one is interested in reducing the emergence of drug resistance, the incidence of HIV, or the occurrence of AIDS. Our model suggests that the preventive benefits from screening depend heavily on the actual proportion of patients who

would truly adhere to HAART. Although reducing the number of drug-resistant HIV infections over 20 years could be achieved with relatively low screening accuracy, the threshold screening accuracy for preventing AIDS cases is quite high at all levels of adherence in the population. The threshold accuracy required to prevent HIV cases is prohibitively high across all scenarios considered. Even perfect screening accuracy would lead to more AIDS cases and HIV infections than no screening over time periods of 20 years and beyond.

What explains these results? Focusing on Figure 4, consider first the goal of avoiding drug resistance. Treating adherent patients increases the likelihood that resistant virus will emerge by a small amount, although treating nonadherent patients greatly increases the emergence of drug-resistant virus. Whether a patient would adhere to therapy or not, withholding treatment does not contribute to the emergence of resistance. If all patients are treated, resistant virus emerges at some average rate in the population depending on the true level of patient adherence. Screening for adherence may result in withholding therapy from those thought to be unable to adhere, which explains why even screening at low levels of accuracy seems attractive at all (relative to the strategy of always offering treatment). The greater the advantage of

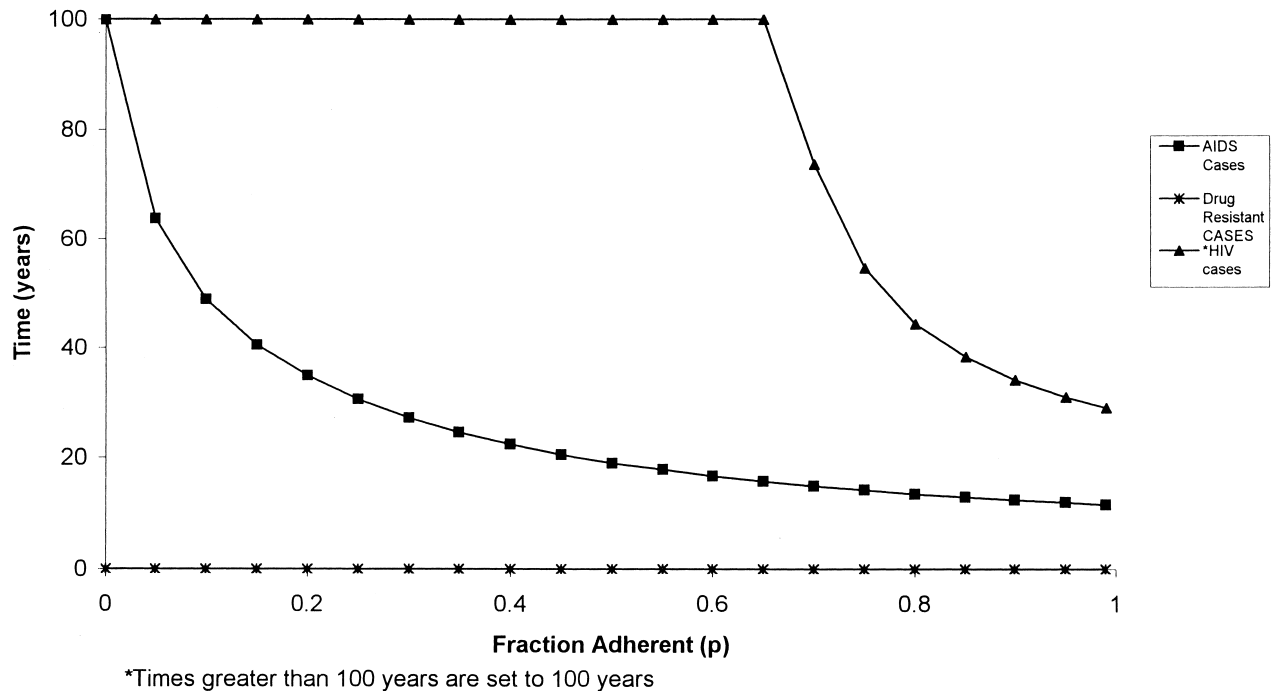


FIG. 5. Time until the total numbers of drug-resistant infections, HIV infections, and AIDS cases are lower with screening for adherence than without.

adherence over nonadherence with respect to the emergence of resistant virus, however, the more serious are the consequences of screening errors. Thus, the accuracy threshold beyond which screening is sensible is most stringent for the “Adherence strongly favorable” scenario, least stringent for the “Adherence weakly stringent” scenario, and intermediate for the “Adherence favorable” scenario. Note that in the extreme, if the goal of AIDS care was simply to prevent the emergence of drug resistance, HAART would never be offered.

Now focus on the goal of preventing AIDS cases. Whether adherent or not and whether infected with drug-sensitive or drug-resistant virus, HAART, on average, does slow progression to AIDS. Although adherent patients fare better than nonadherent patients, nonadherent treated patients fare slightly better than untreated patients. Again, viewing screening for adherence as a mechanism for withholding therapy, at the population level, it now becomes crucial for screening decisions to be accurate. If therapy is withheld from nonadherent patients, it must be that therapy is offered to adherent patients so that, on balance, progression to AIDS in the population is reduced. This, however, requires accurate screening. If the differential between adherent and nonadherent patients is not that great (as in the “Adherence weakly favorable AIDS” scenario in Fig. 4), withholding therapy from adherent patients is less costly than if there

are major benefits from adherence (as in the “Adherence strongly favorable AIDS” scenario). This explains why the screening threshold is more stringent for the latter scenario than for the former when preventing AIDS cases is the goal.

It therefore seems that from the perspective of HIV and AIDS prevention, screening for adherence, even if useful for individual patients, is difficult to justify. The potential effectiveness of HAART as an HIV and AIDS preventive measure could seriously be compromised by such a practice. We have argued that if the screening process is noninformative, as was apparently the case in the study of Paterson et al. (15), screening for adherence is similar to a method of randomly selecting patients to receive treatment. Even with an informative screening process, however, reducing the emergence and transmission of resistant virus may, paradoxically, be outweighed by increases in the number of HIV and AIDS cases over time.

Clearly, if screening for adherence is to be pursued, it is important to develop highly accurate screening methods. Unfortunately, there has been little attention devoted to the public health consequences of screening for adherence. What is truly needed, however, are the commitment and resources to develop the infrastructure necessary to support better adherence among all patients. Because adherence to HAART is the responsibility of both

clinicians and patients, improving adherence should be facilitated by patient-focused and provider-based interventions (22). Developing individualized effective adherence and treatment plans de-emphasizes predicting nonadherent behavior in favor of maximizing the benefits of treatment. We believe this strikes a better balance between concerns regarding the emergence of drug resistance, overall rates of HIV and AIDS in the population, and patients' rights to HAART.

Acknowledgments: Professor Kaplan's research was supported in part by the Societal Institute for the Mathematical Sciences through grant DA-90531 from the National Institute of Drug Abuse, and by the Center for Interdisciplinary Research on AIDS at Yale University through grant MH/DA-56286 from the National Institutes of Mental Health and Drug Abuse.

REFERENCES

1. Deeks SG, Smith M, Holodniy M, et al. HIV-1 protease inhibitors. A review for clinicians. *JAMA* 1997;277:145-53.
2. Condra JH, Schleif WA, Blahy OM, et al. In vivo emergence of HIV-1 variants resistant to multiple protease inhibitors. *Nature* 1995;374:569-71.
3. Little SJ, Daar ES, D'Aquila RT, et al. Reduced antiretroviral drug susceptibility among patients with primary HIV infection [comments]. *JAMA* 1999;282:1142-9.
4. Mayers DL. Drug resistant HIV-1: the virus strikes back. *JAMA* 1998;279:2000-2.
5. Boden D, Hurley A, Zhang L, et al. HIV-1 drug resistance in newly infected individuals. *JAMA* 1999;282:1135-41.
6. Wainberg MA, Friedland GH. Public health implications of antiretroviral therapy and HIV drug resistance. *JAMA* 1998;279:1977-83.
7. Bangsberg DR, Hecht FM, Charlebois ED, et al. Adherence to protease inhibitors, HIV-1 viral load, and development of drug resistance in an indigent population. *AIDS* 2000;14:357-66.
8. Yerly S, Kaiser L, Race E, et al. Transmission of antiretroviral-drug-resistant HIV-1 variants. *Lancet* 1999;354:729-33.
9. Zaric G, Brandeau M, Bayoumi A, et al. The effects of protease inhibitors on the spread of HIV and the development of drug-resistant HIV strains. *Simulation* 1998;74:262-75.
10. Blower SM, Gershengorn HB, Grant RM. A tale of two futures: HIV and antiretroviral therapy in San Francisco. *Science* 2000; 287:650-4.
11. Carpenter C, Cooper DA, Fishl MA, et al. Antiretroviral therapy in adults. Updated recommendations of the International AIDS Society-U.S.A. Panel. *JAMA* 2000;283:381-9.
12. Panel on Clinical Practices for Treatment for HIV Infection. *Guidelines for the use of antiretroviral agents in HIV-infected adults and adolescents*. Washington, DC: Department of Health and Human Services, 2000.
13. Kaplan EH. To treat or not to treat? *Nat Med* 1998;4:664-5.
14. Quinn TC, Wawer MJ, Sewankambo N, et al. Viral load and heterosexual transmission of human immunodeficiency virus type 1. Rakai Project Study Group. *N Engl J Med* 2000;342:921-9.
15. Paterson DL, Swindells S, Mohr J, et al. Adherence to protease inhibitor therapy and outcomes in patients with HIV infection. *Ann Intern Med* 2000;133:21-30.
16. Zhang L, Ramratnam B, Tenner-Racz K, et al. Quantifying residual HIV-1 replication in patients receiving combination antiretroviral therapy. *N Engl J Med* 1999;340:1605-13.
17. Land S, McGavin C, Lucas R, et al. Incidence of zidovudine-

- resistant human immunodeficiency virus isolated from patients before, during, and after therapy. *J Infect Dis* 1992;166:1139-42.
18. Singh N, Squier C, Sivek C, et al. Determinants of compliance with antiretroviral therapy in patients with human immunodeficiency virus: prospective assessment with implications for enhancing compliance. *AIDS Care* 1996;8:261-9.
19. Gordillo V, del Amo J, Soriano V, et al. Sociodemographic and psychological variables influencing adherence to antiretroviral therapy. *AIDS* 1999;13:1763-9.
20. Belzer ME, Fuchs DN, Luftman GS, et al. Antiretroviral adherence issues among HIV-positive adolescents and young adults. *J Adolesc Health* 1999;25:316-9.
21. Friedland GH, Williams AB. Attaining higher goals in HIV treatment. The central importance of adherence. *AIDS* 1999;13(Suppl 1):S61-72.
22. Williams A, Friedland GH. Adherence, compliance, and HAART. *AIDS Clin Care* 1997;9:51-4,58.

APPENDIX

Based on our assumption of symmetric accuracy when screening for adherence, only patients who screen "positive for adherence" receive treatment. Patients in the holding state are screened for adherence at rate Ω . Thus, patients in the holding state either proceed to HAART and adhere (with rate $\Omega p \alpha$) or proceed to HAART but do not adhere (with rate $\Omega(1-p)(1-\alpha)$), where p is the true level of adherence in the holding state patient population, and α is the accuracy of screening for adherence. We assume that whenever a patient enters the holding state, whatever clinician/patient interactions take place lead a patient to adhere with probability p .

For susceptible men, the probability of becoming infected with HIV per sexual partner for drug-sensitive and drug-resistant strains at time t is given by:

$$\lambda_s(t) = \frac{\beta_S^U (Y_S + H_S) + \beta_S^T (Z_S^A + Z_X^{NA}) + k^S \beta_S^U (Y_R + H_R) + k^S \beta_S^T (Z_R^A + Z_R^{NA})}{N(t)}$$

$$\lambda_R(t) = \frac{\beta_R^U (Y_R + H_R) + \beta_R^T (Z_R^A + Z_R^{NA})}{N(t)}$$

where

$$N(t) = X(t) + Y_S(t) + Y_R(t) + H_S(t) + H_R(t) + Z_S^A(t) + Z_S^{NA}(t) + Z_R^A(t) + Z_R^{NA}(t)$$

is the size of the population at time t , and k^S is a propagation factor that modulates the probability that a drug-resistant HIV-infected person transmits a drug-sensitive strain of the virus to his sexual partner as employed by Blower et al. (10). The B 's are HIV transmission probabilities that depend on the treatment status and viral strain of the infected individual.

Table 2 summarizes the state variables of our model, and Figure 1

TABLE 2. State variables in the epidemic model

$X(t)$	Susceptible
$Y_S(t)$	Undetected/drug sensitive
$Y_R(t)$	Undetected/drug resistant
$H_S(t)$	Holding/drug sensitive
$H_R(t)$	Holding/drug resistant
$Z_S^A(t)$	Treated and adherent/drug sensitive
$Z_R^A(t)$	Treated and adherent/drug resistant
$Z_S^{NA}(t)$	Treated and nonadherent/drug sensitive
$Z_R^{NA}(t)$	Treated and nonadherent/drug resistant

presents a flow diagram describing the transition dynamics that relate these state variables. The model is defined mathematically by the following equations:

$$\begin{aligned} \frac{dX(t)}{dt} &= \prod - \{c(\lambda_S(t) + \lambda_R(t)) + \mu\} X(t) \\ \frac{dY_S(t)}{dt} &= c\lambda_S X(t) + qY_R(t) - (\mu + \gamma_S^U + \tau)Y_S(t) \\ \frac{dY_R(t)}{dt} &= c\lambda_R X(t) + (q + \mu + \gamma_R^U + \tau)Y_R(t) \\ \frac{dH_S(t)}{dt} &= \tau Y_S(t) + qH_R(t) + g_S^A Z_S^A(t) + g_S^{NA} Z_S^{NA}(t) \\ &\quad - \{\mu + \gamma_S^U + \Omega[\alpha p + (1-p)(1-\alpha)]\} H_S(t) \\ \frac{dH_R(t)}{dt} &= \tau Y_R(t) - qH_R(t) + \theta^A Z_S^A(t) + \theta^{NA} Z_S^{NA}(t) + g_R^A Z_R^A(t) + g_R^{NA} Z_R^{NA}(t) \\ &\quad - \{\mu + \gamma_R^U + \Omega[\alpha p + (1-p)(1-\alpha)]\} H_R(t) \\ \frac{dZ_S^A(t)}{dt} &= \Omega p \alpha H_S(t) - (\mu + \gamma_S^A + g_S^A + \theta^A) Z_S^A(t) \\ \frac{dZ_S^{NA}(t)}{dt} &= \Omega(1-p)(1-\alpha) H_S(t) - (\mu + \gamma_S^{NA} + g_S^{NA} + \theta^{NA}) Z_S^{NA}(t) \\ \frac{dZ_R^A(t)}{dt} &= \Omega p \alpha H_R(t) - (\mu + \gamma_R^A + g_R^A) Z_R^A(t) \\ \frac{dZ_R^{NA}(t)}{dt} &= \Omega(1-p)(1-\alpha) H_R(t) - (\mu + \gamma_R^{NA} + g_R^{NA}) Z_R^{NA}(t) \end{aligned}$$

Table 1 summarizes parameter values for the “Adherence favorable,” “Adherence weakly favorable,” and “Adherence strongly favorable” scenarios described in the text. To obtain these parameter estimates, we employed published estimates of progression rates, mutation rates, rates of withdrawal from therapy, and the proportion of currently treated patients believed to adhere to HAART (60%). Consider the rate of progressing from HIV to AIDS for those infected with the sensitive strain of the virus. Defining κ as the progression rate penalty associated

with nonadherence to HAART, the following equation determines the progression rate, γ_S^A , for adherent patients infected with the drug-sensitive strain:

$$0.6\gamma_S^A + 0.4\gamma_S^{NA} = 0.6\gamma_S^A + 0.4 \times \kappa \times \gamma_S^A = 0.037$$

where 0.037 per year is the mean progression rate for drug-sensitive patients on HAART obtained from the probability distribution described by Blower and colleagues (10). The progression rate for non-adherent HAART patients infected with the sensitive strain, γ_S^{NA} , is then set equal to $\kappa \times \gamma_S^A$. Given the value of κ and similar mean parameter values for drug-resistant mutation rates and rates of withdrawal from therapy as taken from Blower and colleagues’ probability distributions (10), the estimates summarized in Table 1 are easily obtained from formulas that correspond to that described previously. The penalty parameters (κ ’s) used to derive these different scenarios are similar to values suggested by Zaric et al. (9) and are shown in Table 3.

We also model scenarios where all patients receive HAART in the absence of screening for adherence and scenarios where no therapy is offered. When HAART is offered in the absence of screening for adherence, the model is modified by interpreting the parameter Ω as the per patient rate at which individuals leave the holding state and enter therapy (or equivalently, all screening decisions would lead to treatment). Any patient newly receiving therapy adheres to HAART with probability p and fails to adhere with probability $1 - p$. When no therapy is offered, the parameter Ω is simply set equal to zero, which implies that the state variables $Z_S^A(t)$, $Z_S^{NA}(t)$, $Z_R^A(t)$, and $Z_R^{NA}(t)$ which keep track of the number of treated individuals, equal zero as well.

TABLE 3. Progression rate penalty factors (κ ’s)

	Adherence favorable	Adherence weakly favorable	Adherence strongly favorable
Rate of progression	1.5	1.4	2.0
Rate of mutation	10.0	5.0	80.0
Rate of withdrawal	3.5	2.0	200.0