THE IMPACT OF MORTALITY ON PANJER'S MODEL OF AIDS SURVIVAL

COLIN M. RAMSAY

ABSTRACT

This paper is intended to be a sequel to Panjer's paper [12]. Using Ramsay's [13] extension to Panjer's model, we prove that Panjer's results are biased. In particular, his estimates of the transition intensities are too high because he did not include the mortality risk that exists before a life develops full-blown AIDS. Panjer thus understates the average time that it will take an HIV+ life to develop AIDS or to die of AIDS. Suggestions are made for gathering data in an AIDS environment.

1. INTRODUCTION

The Walter Reed Staging Method (WRSM), as described by Redfield et al. [14], is a method that groups patients infected with the Human Immunodeficiency Virus (HIV) into various stages, the final stage being AIDS. The WRSM was used by Panjer [12] to describe the progression of the disease process once a life became infected with HIV. His model essentially says that the progression of the disease through its various stages is sequential and irreversible. The stages are:

Stage 0 (At-risk) Healthy persons at risk for HIV infection, but testing negative
Stage 1 (HIV+) Otherwise asymptomatic persons testing HIV+
Stage 2 (LAS) Persons with HIV infection and lymphademopathy syndrome (LAS), together with moderate celluar immune deficiency
Stage 3 (ARC) Patients with HIV infection and LAS, together with severe cellular immune deficiency (AIDS-Related Complex, or ARC)
Stage 4 (AIDS) Patients with acquired immune deficiency syndrome (AIDS)
Stage 5 (Death) Patients who died "of AIDS."
....
Stage 6 (Death) Persons who died in stages 0, 1, 2, or 3. Persons who died in stage 4 of causes not related to AIDS.

In stage 0 are healthy lives that are not currently testing positive for HIV but have a relatively high probability of becoming infected with the virus. Such a group of lives is called an "at-risk" group. Several factors contribute to exposing a life to the risk of becoming HIV+. The factors cited most often are IV drug use, sexual orientation and sexual promiscuity.

To avoid confusion, we must define the conditions under which the term "AIDS" is used. *In this paper a life is said to have AIDS if and only if it is in stage 4.* This definition is in contrast to the popular usage that refers to any life in stages 1 through 3 as having "AIDS." Obviously in these earlier stages the lives have not yet developed AIDS in the sense of the WRSM, so the term "AIDS" is not used to describe such lives. As a result, the expression "a life died of AIDS" means that the life died in stage 4 from a disease or condition that is directly related to the AIDS condition.

Clearly a life can die at any time and in any of the stages. This fact was not included in Panjer's model. If one assumes that the mortality pattern exhibited by the members of an at-risk group is similar to standard mortality, one may be able to argue that the force of mortality will be negligible when compared to the force of transition to the next stage according to the WRSM. However, if one thinks that lives in the high-risk groups have significantly substandard mortality experiences, it will be necessary to include a mortality component in the model at stage 0. Once a person has entered stage 1, it is clear that the person's overall mortality and morbidity rates will be worse than those of healthy uninfected lives, especially as the life progresses through the various stages to AIDS. This is because HIV attacks the body's T4 lymphocyte cells, thus destabilizing the body's immune system and leaving the body open to opportunistic infections. Hence in stages 1 through 3 the death rates from life-threatening diseases will be expected to be significantly higher in HIV+ populations than in uninfected ones. In view of this, the effects of mortality on lives *before* they have developed AIDS cannot be ignored.

The inclusion of the possibility of death in any stage and the separation of the deaths in stage 4 into AIDS- and non-AIDS-related causes are not new ideas. This was the approach taken by the Institute of Actuaries AIDS Working Party in its reports (see Daykin [3, pp. 142–144]) and by most statisticians and epidemiologists developing mathematical models of the dynamics of HIV/AIDS (see, for example, Dietz [4], Hyman and Stanley [7], Isham [8], and May [11]). In view of this, we are adding a seventh stage, labeled stage 6, to accommodate the earlier deaths. Hence death in any of stages 0 through 3 means immediate transition to stage 6. In stage 4 the term "death" is reserved for death due to causes that are unrelated to AIDS and its complications; these deaths imply immediate transition to stage 6. A person who has died of AIDS or its complications is said to have "progressed" to stage 5.

2. THE MODEL WITH MORTALITY

Not much work has been done in the area of mathematically modeling the *transition* process of HIV by using the WRSM. Simple models of the transition dynamics of HIV have been developed by Cowell and Hoskins [2], Panjer [12], and Ramsay [13]. Cowell and Hoskins used an essentially nonparametric, discrete time model of the probabilities of progression. They also included "duration since progression" as a factor in the transition probabilities. However, as Panjer pointed out, their results are highly sensitive to their data. On the other hand, Panjer developed a model of an "at-risk life" in stage 0 and its progression through to AIDS and then death. Death prior to the development of full-blown AIDS was not permitted in this model. He used a parametric, continuous time Markov process model with constant forces of transition from one stage to the next. Ramsay extended Panjer's model to include a death term. He then developed expressions for life insurance functions.

Following Ramsay, we adapt Panjer's model to include a death component as follows: A person in stage i ($i = 0, 1, 2, 3, 4$) is assumed to be subject to a constant force of progression (μ_i) out of stage i and into stage $i+1$, and a constant force of mortality (μ_i') out of stage i and directly into stage 6. These forces are assumed to be operating simultaneously on each person. In view of this, one must clearly distinguish between these forces for patients with AIDS because they can die from AIDS-related diseases or from non-AIDS-related diseases. In stage 4, the former type of death is considered to be due to μ_4, while the latter is due to μ_4'.

The assumption of constant forces in each stage may appear to be somewhat unrealistic. However, this assumption is used quite often when developing mathematical models of HIV/AIDS transmissions (see, for example, Dietz [4], Hyman and Stanley [7], Isham [9], and May [11]). Panjer's results showed that this is a plausable first approximation to the actual transition process. In keeping with the famous Occam's Razor—hypotheses should not be complicated unnecessarily—we made the simplest and least complicated extension to Panjer's model.

Because the forces are assumed to be constant, a "memoryless" property will exist. This means that the length of time already spent in the current stage has no effect on the future length of time that the person will remain in this stage. This fact will allow us to speak in terms of the future time spent in a stage without having to condition on the amount of time already spent in the stage.

Definition 1. Let T_i be the (future) time spent in stage i before entering stage $i+1$, and T_i' be the (future) time spent in stage i before immediate transition to stage 6, that is, death in stage i. Due to the "memoryless" property described above, the random variables $\{T_0, \ldots, T_{i-1}, T_i'\}$ are mutually independent as are the random variables.

Let $f_i(t)$ and $g_i(t)$ be the probability density functions (pdf's) of T_i and T_i', respectively; then $f_i(t)$ and $g_i(t)$ are given by

$$f_i(t) = \left(\frac{\mu_i}{\alpha_i}\right)\alpha_i e^{-\alpha_i t} \tag{1}$$

$$g_i(t) = \left(\frac{\mu_i'}{\alpha_i}\right)\alpha_i e^{-\alpha_i t} \tag{2}$$

where α_i is the force of transition out of stage i, but not necessarily into stage $i+1$, that is,

$$\alpha_i = \mu_i + \mu_i'. \tag{3}$$

At this point note that, technically, neither $f_i(t)$ nor $g_i(t)$ is a pdf because neither of them integrates to 1. As a consequence, one cannot regard T_i and T_i' as completely specified random variables because some part of their "randomness" is missing. In view of this, the following extensions will be made to the definition of T_i and T_i': (1) if the life progresses to stage $i+1$, then we say that $T_i' = \infty$, and (2) if the life dies while in stage i, then $T_i = \infty$. As a result of the extended definition of T_i and T_i', their moments will not exist; that is, they will be infinite. However, the life expectancy of a life in stage i can still be evaluated because, among other things, death (from any cause) is certain, so random variables such as the total future lifetime until death (from any cause) for a life currently in stage i will be well defined with finite moments.

3. THE BIAS IN PANJER'S ESTIMATES

Because Panjer's model did not include a mortality component in stages 0 through 3, it can be viewed as one based on lives that are destined (or predetermined) to die of AIDS. His model is called a "conditional model" because it is equivalent to conditioning on the event that all the lives will eventually die of AIDS. In the conditional model the random variable T_i' will obviously not exist because such deaths are not permitted.

Let $f_i^c(t)$ be the pdf of the random variable T_i conditioned on the event that the life died of AIDS. For notational convenience, define the event $\mathcal{A}_i$ as

$$\mathcal{A}_i = \{\text{the life currently in stage } i \text{ will die of AIDS}\}$$

for $i = 0, 1, 2, 3, 4$. Now $f_i^c(t)$ can be found by using the definition of conditional probability as applied to the WRSM.

$$f_i^c (t)\, dt = \Pr[t < T_i \le t + dt | \mathcal{A}_i]$$

$$= \frac{\Pr[t < T_i \le t + dt \cap \mathcal{A}_i]}{\Pr[\mathcal{A}_i]}$$

$$= \frac{\Pr[t < T_i \le t + dt]\, \Pr[\mathcal{A}_{i+1}]}{\Pr[\mathcal{A}_i]}$$

By the definition of $\mathcal{A}_i$ and Equation (1),

$$\Pr[\mathcal{A}_i] = \int_0^\infty f_i(t)\, \Pr[\mathcal{A}_{i+1}]\, dt$$

$$= \frac{\mu_i}{\alpha_i}\, \Pr[\mathcal{A}_{i+1}].$$

Therefore, the pdf of the random variable $\{T_i | \mathcal{A}_i\}$ is

$$f_i^c(t) = (f_i(t)\, \Pr[\mathcal{A}_{i+1}]) / \left(\frac{\mu_i}{\alpha_i}\, \Pr[\mathcal{A}_{i+1}] \right)$$

$$= \alpha_i\, e^{-\alpha_i t}, \tag{4}$$

an exponential distribution with parameter α_i.

The exponential distribution with parameter μ_i was the pdf of Panjer's version of T_i; see Panjer's Equation (3). Therefore, Panjer's version of T_i is actually the random variable $\{T_i | \mathcal{A}_i\}$. This means that Panjer in fact estimated α_i instead of μ_i, making his estimates of the μ_i's biased; he is overstating the size of each μ_i. As a consequence, the expected time that it will take an HIV+ life to develop AIDS will be longer than the expected time given by Panjer. In fact, in a strictly mathematical sense, this expected time will not even exist. Techniques for deriving maximum likelihood estimates of the parameters μ_i and μ_i' are provided in the Appendix.

4. SUMMARY AND COMMENTS

Panjer's parametric survival model for HIV+ lives was extended to include mortality prior to the development of full-blown AIDS. It was proved that Panjer's model is a conditional model with an intensity function in stage i equal to α_i, the sum of the forces of mortality and progression to stage $i+1$. This produced a bias in Panjer's estimates of the μ_i's, leading to an overestimation of the μ_i's. This meant that the expected time it takes an HIV+ to develop AIDS will be longer than the expected time given by Panjer. To estimate the force of progression without using mortality, one must observe the lives in the study without the effects of mortality before stage 4. Unfortunately, this is not possible. As a result, the questions posed by Panjer—"What is the probability that an HIV+ person will develop AIDS within the next 3 years?" or "Of 1000 HIV+ persons, how many can we expect to have AIDS or have died of AIDS within the next 5 years?"— cannot be answered with his model. However, by using Ramsay's model and his expressions for the various transition probabilities, these probabilities can now easily be found.

One of the most important ingredients in the survival analysis of HIV+ lives is the data. The data used by Panjer did not contain the information necessary to jointly estimate the mortality and transition parameters for even the simple model described in this paper. Clearly, if we want to introduce a more complex model, for example, one that assumes the intensity functions are dependent on the duration since entering a stage, more sophisticated studies and systems of data collection will be necessary. Future studies based on the WRSM must, at a minimum, generate the following five pieces of information to a tolerable degree of accuracy: (1) date of initial infection, (2) date of transition to each stage, (3) date of death in study or date of exit from study, (4) stage in which death occurs, and (5) cause of death, that is, AIDS-related or non-AIDS-related cause of death.

The fifth piece of information is most likely to be found on a death certificate. This immediately raises two concerns: (i) What constitutes an AIDS-related death? and (ii) Are such deaths reliably reported? The former concern can be dealt with by using the Centers for Disease Control's (CDC) definition of AIDS. The CDC [15] defined AIDS in terms of a set of "indicator" diseases; deaths that result from any of these diseases are termed AIDS-related deaths. As for the second concern, it may seem reasonable to expect the death certificates of HIV/AIDS patients to be frequently unreliable because attendant physicians may be reluctant to disclose the true cause of

death in such cases. However, the recent work by Hardy et al. [6] showed that this may not be the case. After reviewing death certificates in four U.S. cities—Washington, D.C., New York City, Boston, and Chicago—to evaluate AIDS surveillance effectiveness, Hardy et al. found nearly complete AIDS case reporting. The level of reporting was much higher than reporting for many other communicable diseases. For insurance purposes, the AIDS Task Force [17, chapter 4, p. 11] suggested that a working definition of a "suspected AIDS" death needs to be developed to help offset the veiled and misstated cause-of-death problem. They provided 11 AIDS-related diseases that may be helpful in developing such a definition.

In closing, the number of persons developing full-blown AIDS is, metaphorically speaking, just the tip of the iceberg. Attention must be directed to the totality of HIV-infected lives in the population. Rolland [16, p. 6] wrote "Our enemy is HIV infection, not AIDS per se. Given that a very high percentage of those who become infected will die, one's focus must be on all who are infected rather than on that small subset who have progressed all the way to AIDS."

It is hoped that studies on the mortality, morbidity and transition processes of lives in the various stages along the progression to AIDS will be undertaken, and that data will be gathered in a manner that will facilitate a parametic analysis of these processes.

REFERENCES

1. BURDEN, R.L. AND FAIRES, J.D. *Numerical Analysis,* 3rd ed. Boston: Pringle, Weber & Schmidt, 1985.
2. COWELL, M.J. AND HOSKINS, W.H. "AIDS, HIV Mortality and Life Insurance," Chapter 3 in *The Impact of Aids on Life and Health Insurance Companies: A Guide for Practicing Actuaries, Report of the Society of Actuaries Task Force on AIDS.* In *TSA* XL, Part II (1988): 909–72.
3. DAYKIN, C.D. "A United Kingdom Perspective." In *Insurance and the AIDS Epidemic, Proceedings of a Two Day Symposium.* Schaumburg, Ill.: Society of Actuaries, 1988, pp. 133–167.
4. DIETZ, K. "On the Transmission Dynamics of AIDS," *Mathematical Biosciences* 90 (1988): 397–414.
5. ELANDT-JOHNSON, R.C. AND JOHNSON, N.L. *Survival Models and Data Analysis.* New York: Wiley, 1980.
6. HARDY, A.M., STARCHER, E.T., ET AL. "Review of Death Certificates to Assess Completeness of AIDS Case Reporting," *Public Health Report* 102 (1987): 386–91.

7. HYMAN, J.M. AND STANLEY, E.A. "Using Mathematical Models to Understand the AIDS Epidemic," *Mathematical Biosciences* 90 (1988): 415–73.

8. IMAGAWA, D.T., ET AL. "Human Immunodeficiency Virus Type 1 Infection in Homosexual Men Who Remain Seronegative for Prolonged Periods," *New England Journal of Medicine* 320, no. 22 (June 1, 1989): 1458–62.

9. ISHAM, V. "Mathematical Modelling of the Transmission Dynamics of HIV Infection and AIDS: A Review," *Journal of the Royal Statistical Society* Ser. A, Part 1 (1988): 5–30.

10. KENDALL, SIR M.G. AND STUART, A. *The Advanced Theory of Statistics,* Vol. 2, 4th ed. New York: Macmillan, 1979.

11. MAY, R.M., ANDERSON, R.M AND MCLEAN, A.R. "Possible Demographic Consequences of HIV/AIDS Epidemics. 1. Assuming HIV Infection Always Leads to AIDS," *Mathematical Biosciences* 90 (1988): 475–505.

12. PANJER, H.H. "AIDS: Survival Analysis of Persons Testing HIV +," *TSA* XL (1988): 517–30.

13. RAMSAY, C.M. "AIDS and the Calculation of Life Insurance Functions," *TSA* XLI (1989): 393–422.

14. REDFIELD, R., WRIGHT, D.C. AND TRAMONT, E.C. "The Walter Reed Staging Classification for HIV-III / LAV Infection," *New England Journal of Medicine* 314, no. 2 (1986): 131–32.

15. "Revision of the CDC Surveillance Case Definition for Acquired Immunodeficiency Syndrome," *Centers for Disease Control Morbidity and Mortality Weekly Supplement* 36, no. 1S (1987).

16. ROLLAND, I.M. "The Impact of AIDS on American and Canadian Insurers." In *Insurance and the AIDS Epidemic, Proceedings of a Two Day Symposium.* Schaumburg, Ill.: Society of Actuaries, 1988, pp. 1–21.

17. SOCIETY OF ACTUARIES AIDS TASK FORCE. "The Impact of AIDS on Life and Health Insurance Companies: A Guide for Practicing Actuaries," *TSA* XL, Part II (1988): 839–1160.

APPENDIX
ESTIMATION OF PARAMETERS

Let us consider a study consisting of m observation periods, the j-th period being $(a_j, b_j), j = 1, 2, \ldots, m$. A group of lives n_{ij} in some stage arbitrarily labelled i is followed from the *start* of the j-th period until the *end* of this period. During this period lives can remain in stage i, progress to stage $i + 1$ and beyond, or die while in stage i. The number of persons dying while in stage i and the number of persons progressing to stage $i + 1$ and beyond are recorded separately. The information gathered from each of the m periods will be used to estimate the forces of mortality and progression.

For $i = 0, 1, 2, 3, 4$ and $j = 1, \ldots, m$, define

μ_i = force of progression from stage i to $i + 1$

μ_i' = force of progression from stage i to 6

$\alpha_i = \mu_i + \mu_i'$

d_{ij} = number of lives progressing from stage i to $i + 1$ during observation period j

d_{ij}' = number of lives progressing from stage i to stage 6 during observation period j

$\bar{p}_{ij}$ = probability that a life in stage i at the start of period j will remain in stage i throughout this period

$\bar{q}_{ij}$ = probability that a life in stage i at the start of period j will enter stage $i + 1$ during period j

$\bar{q}_{ij}'$ = probability that a life in stage i will progress directly to stage 6.

Because of the "memoryless" property, the probabilities defined above will depend only on the length of period j and the current stage i. The subscript i will be dropped from all symbols. Clearly the probabilities introduced above are given by

$$\bar{p}_j = \int_{(b_j - a_j)}^{\infty} \alpha e^{-\alpha t}\, dt$$

$$= e^{-\alpha(b_j - a_j)} \tag{5}$$

$$\bar{q}_j = \int_0^{(b_j - a_j)} \mu e^{-\alpha t}\, dt$$

$$= \frac{\mu}{\alpha}\left(1 - e^{-\alpha(b_j - a_j)}\right) \tag{6}$$

$$\bar{q}'_j = \int_0^{(b_j - a_j)} \mu' e^{-\alpha t}\, dt$$

$$= \frac{\mu'}{\alpha}\left(1 - e^{-\alpha(b_j - a_j)}\right) \tag{7}$$

with the subscripts i dropped.

From Elandt-Johnson and Johnson [5, chapter 12.2.1, p. 324, equation 12.4], the likelihood L is the product of m multinomial distributions, that is,

$$L(\mu, \mu') = \prod_{j=1}^{m} \binom{n_j}{d'_j \ (n_j - d_j - d'_j) \ d_j} \bar{q}_j^{\,d_j} \bar{q}_j'^{\,d'_j} \bar{p}_j^{\,(n_j - d_j - d'_j)}$$

while the log-likelihood is given by

$$l(\mu, \mu') = const + \sum_{j=1}^{m} [(n_j - d_j - d'_j) \ log \ \bar{p}_j$$
$$+ \ d_j \ log \ \bar{q}_j + d'_j \ log \ \bar{q}_j']. \quad (8)$$

To obtain the maximum likelihood estimators of μ and μ', we must solve the pair of simultaneous equations $\partial l/\partial \mu = 0$ and $\partial l/\partial \mu' = 0$, where

$$\frac{\partial l}{\partial \mu} = \sum_{j=1}^{m} \left[\frac{n_j - d_j - d'_j}{\bar{p}_j} \frac{\partial \bar{p}_j}{\partial \mu} + \frac{d_j}{\bar{q}_j} \frac{\partial \bar{q}_j}{\partial \mu} + \frac{d'_j}{\bar{q}_j'} \frac{\partial \bar{q}_j'}{\partial \mu} \right] \quad (9)$$

and

$$\frac{\partial l}{\partial \mu'} = \sum_{j=1}^{m} \left[\frac{n_j - d_j - d'_j}{\bar{p}_j} \frac{\partial \bar{p}_j}{\partial \mu'} + \frac{d_j}{\bar{q}_j} \frac{\partial \bar{q}_j}{\partial \mu'} + \frac{d'_j}{\bar{q}_j'} \frac{\partial \bar{q}'_j}{\partial \mu'} \right]. \quad (10)$$

From Equations (25) to (27), it is seen that

$$\frac{\partial \bar{p}_j}{\partial \mu} = -(b_j - a_j)\bar{p}_j \quad (11)$$

$$\frac{\partial \bar{p}_j}{\partial \mu'} = -(b_j - a_j)\bar{p}_j \quad (12)$$

$$\frac{\partial \bar{q}_j}{\partial \mu} = \frac{\mu'}{\mu \alpha} \bar{q}_j + \frac{\mu}{\alpha}(b_j - a_j)\bar{p}_j \quad (13)$$

$$\frac{\partial \bar{q}_j}{\partial \mu'} = -\frac{1}{\alpha}\bar{q}_j + \frac{\mu}{\alpha}(b_j - a_j)\bar{p}_j \quad (14)$$

$$\frac{\partial \bar{q}_j'}{\partial \mu} = \frac{\mu}{\mu' \alpha} \bar{q}_j + \frac{\mu'}{\alpha}(b_j - a_j)\bar{p}_j \quad (15)$$

$$\frac{\partial \bar{q}_j'}{\partial \mu'} = -\frac{1}{\alpha}\bar{q}_j + \frac{\mu'}{\alpha}(b_j - a_j)\bar{p}_j \quad (16)$$

Substituting Equations (11) to (16) into Equations (9) and (10) will yield the desired pair of simultaneous equations for finding the maximum likelihood estimates of μ and μ'. One can use the multivariate version of the Newton-Raphson method described in Burden and Faires [1, chapter 9.2] to obtain these estimates as follows: Let

$$\underline{\mu} = \begin{pmatrix} \mu \\ \mu' \end{pmatrix}$$

and

$$J(\underline{\mu}) = \begin{bmatrix} \dfrac{\partial^2 l}{\partial^2 \mu} & \dfrac{\partial^2 l}{\partial \mu \partial \mu'} \\[2ex] \dfrac{\partial^2 l}{\partial \mu' \partial \mu} & \dfrac{\partial^2 l}{\partial^2 \mu'} \end{bmatrix}.$$

In order to solve the equation

$$\begin{pmatrix} \dfrac{\partial l}{\partial \mu} \\[2ex] \dfrac{\partial l}{\partial \mu'} \end{pmatrix} = 0,$$

one must choose a starting value $\underline{\mu}^{(0)}$, then for $k = 0, 1, \ldots$ define

$$\underline{\mu}^{(k+1)} = \underline{\mu}^{(k)} - [J(\underline{\mu}^{(k)})]^{-1} \begin{pmatrix} \dfrac{\partial l}{\partial \mu} \\[2ex] \dfrac{\partial l}{\partial \mu'} \end{pmatrix}_{\underline{\mu} = \underline{\mu}^{(k)}}. \tag{17}$$

An intuitive starting value can be found by using the estimator given by Elandt-Johnson and Johnson [5, chapter 12.2.1, p. 325, equation 12.5], that is,

$$\mu^{(0)} = -\frac{1}{m} \sum_{j=1}^{m} \left[\frac{d_j}{(d'_j + d_j)} \frac{1}{(b_j - a_j)} \log \left(\frac{n_j - d_j - d'_j}{n_j} \right) \right] \quad (18)$$

$$\mu'^{(0)} = -\frac{1}{m} \sum_{j=1}^{m} \left[\frac{d'_j}{(d'_j + d_j)} \frac{1}{(b_j - a_j)} \log \left(\frac{n_j - d_j - d'_j}{n_j} \right) \right] \quad (19)$$

If $\hat{\mu}$ is the maximum likelihood estimator of μ, it is well-known that $\hat{\mu}$ is asymptotically multivariate normal with mean μ and variance-covariance matrix $-[J(\hat{\mu})]^{-1}$ (see, for example, Kendall and Stuart [10, chapter 18]).

In the analysis described above, it is assumed that the lives are followed throughout the entire period, and the period did not necessarily start from time 0. However, Panjer assumed that lives were observed over some interval $(0, t_j)$ where t_j is a random variable that lies in the range (a_j^*, b_j^*). Two different scenarios were postulated in his paper: (1) lives were observed only up to the midpoint of each range, and (2) there was a uniform censoring mechanism for each observation. His assumptions will affect our calculations only through their effects on $\bar{p}_j$, $\bar{q}_j$, and $\bar{q}'_j$.

1. Midpoint Assumption

Here it is assumed that for period j, t_j is not random but is the midpoint of the range. This leads to

$$t_j = \frac{a_j^* + b_j^*}{2} \quad (20)$$

$$\bar{p}_j = \int_{t_j}^{\infty} e^{-\alpha t} \, dt$$

$$= e^{-\alpha t_j} \quad (21)$$

$$\bar{q}_j = \int_{0}^{t_j} \mu e^{-\alpha t} \, dt$$

$$= \frac{\mu}{\alpha} (1 - e^{-\alpha t_j}) \quad (22)$$

$$\bar{q}'_j = \int_{0}^{t_j} \mu' e^{-\alpha t} \, dt$$

$$= \frac{\mu'}{\alpha} (1 - e^{-\alpha t_j}). \quad (23)$$

2. Uniform Random Censoring Assumption

Here it is assumed that for period j, t_j has a uniform distribution on (α_j^*, b_j^*).

$$\bar{p}_j = \int_{a_j^*}^{b_j^*} \frac{1}{(b_j^* - a_j^*)} \int_{t_j}^{\infty} \alpha e^{-\alpha s}\, ds\, dt_j$$

$$= \frac{e^{-a_j^*\alpha} - e^{-b_j^*\alpha}}{(b_j^* - a_j^*)\,\alpha} \tag{24}$$

$$\bar{q}_j = \int_{a_j^*}^{b_j^*} \frac{1}{(b_j^* - a_j^*)} \int_0^{t_j} \mu e^{-\alpha s}\, ds\, dt_j$$

$$= \frac{\mu}{\alpha}\left(1 - \frac{(e^{-a_j^*\alpha} - e^{-b_j^*\alpha})}{(b_j^* - a_j^*)\alpha}\right) \tag{25}$$

$$\bar{q}_j' = \int_{a_j^*}^{b_j^*} \frac{1}{(b_j^* - a_j^*)} \int_0^{t_j} \mu' e^{-\alpha s}\, ds\, dt_j$$

$$= \frac{\mu'}{\alpha}\left(1 - \frac{(e^{-a^*_{aj}\alpha} - e^{-b_j^*\alpha})}{(b_j^* - a_j^*)\,\alpha}\right) \tag{26}$$

For each assumption, the partial derivatives of the probabilities can be found and substituted into Equations (9) and (10) to yield the maximum likelihood estimates of the parameters.

Unfortunately the data used by Panjer and by Cowell and Hoskins cannot be used to estimate the parameters in this model because there are no data on the number of transitions directly to stage 6. It is hoped that the appropriate data will become available to the author (or to any other interested party) for the purpose of estimating the parameters and testing the fit of this model.
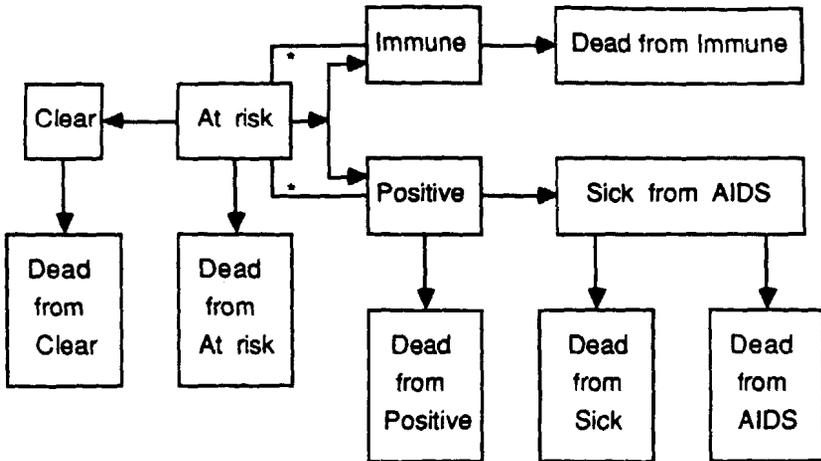
DISCUSSION OF PRECEDING PAPER

ERIC S. SEAH:

I congratulate the author for writing a fine paper on the important topic of AIDS. In this discussion, I would like to briefly describe a model developed in the United Kingdom by Wilkie [1].

In this U.K. model, a life can be in any of the eleven states (see Figure 1). Note that the "Sick from AIDS" state in the U.K. model corresponds to stages 2 to 4 combined in both Panjer's model and Ramsay's model (which is an extension to Panjer's). Deaths during the "Positive" state (which corresponds to stage 1, HIV+, in Panjer's and Ramsay's models) are provided for in the U.K. model. As pointed out by Ramsay, a shortcoming of Panjer's model is that there is no such provision for an HIV+ life before a full-blown AIDS is developed.

FIGURE 1

AIDS MODEL: STATES AND TRANSITIONS

* denotes possible infection.

One useful feature in the U.K. model is that the mortality rates are age-specific, and forces of progression to other states and of remaining in the

current state depend either on age or on both age and duration since entry to the state. For example, the force of progression from the "Positive" state to the "Sick from AIDS" state is a function of both age and the duration since entry to the "Positive" state, and the mortality rate due to AIDS while in the "Sick from AIDS" state is a function of age. Ramsay has noted in the paper that "the assumption of constant forces in each stage may appear to be unrealistic." It would be interesting to see how the Ramsay model can be extended further to take age and duration into account.

REFERENCE

1. WILKIE, A. D. "An Actuarial Model for AIDS," *Journal of the Institute of Actuaries* 115 (1988): 839–53.

BEDA CHAN:

Dr. Ramsay has pointed out again [2] that one may die in stages 0, 1, 2, or 3 and may die in stage 4 but with non-AIDS causes. While we accept the WRSM reasoning that the stages are treated as physiologically sequential, an individual may pass through stages so quickly relative to the frequency of examination and observation that one is considered to have jumped stages. We thus allow jumping to advanced stages as well as jumping directly to the most advanced stage of death. To simplify our model, we do not allow death from non-AIDS causes. Let

$$\lambda^k \mu_i = \text{hazard rate from stage } i \text{ to stage } i + 1 + k$$

where μ_i is the hazard rate from stage i to $i+1$. In other words, we assume that the leaping rate is geometrically scaled down by the distance of the leap, and for the sake of parsimony, we assume the leaping factor λ is the same for all stages. We consider the hazard matrix

$$
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
\mu_0 & 0 & 0 & 0 & 0 & 0 \\
\lambda\mu_0 & \mu_1 & 0 & 0 & 0 & 0 \\
\lambda^2\mu_0 & \lambda\mu_1 & \mu_2 & 0 & 0 & 0 \\
\lambda^3\mu_0 & \lambda^2\mu_1 & \lambda\mu_2 & \mu_3 & 0 & 0 \\
\lambda^4\mu_0 & \lambda^3\mu_1 & \lambda^2\mu_2 & \lambda\mu_3 & \mu_4 & 0
\end{bmatrix}
$$

where λ is the only additional parameter. In particular, the hazard rates to death in our model are:

$$\lambda^4 \mu_0 = \text{hazard rate from stage 0 to death}$$

$$\lambda^3 \mu_1 = \text{hazard rate from stage 1 to death}$$

$$\lambda^2 \mu_2 = \text{hazard rate from stage 2 to death}$$

$$\lambda \mu_3 = \text{hazard rate from stage 3 to death}$$

$$\mu_4 = \text{hazard rate from stage 4 to death,}$$

which are comparable to that of Ramsay's [2, (54)]:

$$\mu_i' = B \, c^i \text{ for } i = 0, 1, 2, 3.$$

Our model here allows for a geometric feature for the hazard rate to death, like Ramsay, but also feature μ_i as a factor that reflects the stability of stage i. The current paper by Ramsay would have

$$
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
\mu_0 & 0 & 0 & 0 & 0 & 0 \\
0 & \mu_1 & 0 & 0 & 0 & 0 \\
0 & 0 & \mu_2 & 0 & 0 & 0 \\
0 & 0 & 0 & \mu_3 & 0 & 0 \\
\mu_0' & \mu_1' & \mu_2' & \mu_3' & \mu_4 & 0
\end{bmatrix}
$$

as its hazard matrix. The hazard matrix for the original Panjer paper is

$$
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
\mu_0 & 0 & 0 & 0 & 0 & 0 \\
0 & \mu_1 & 0 & 0 & 0 & 0 \\
0 & 0 & \mu_2 & 0 & 0 & 0 \\
0 & 0 & 0 & \mu_3 & 0 & 0 \\
0 & 0 & 0 & 0 & \mu_4 & 0
\end{bmatrix}
$$

With the Frankfurt data, this hazard matrix is the best one can do. Our idea of considering jumps towards advanced stages as well as to death is motivated by [1], is which Panjer considers distribution of the incubation period, that is, movement through a number of stages.

We thank the author for pointing out the need for more detail in the data; otherwise the deaths in stages 0, 1, 2, or 3 and deaths in stage 4 with non-AIDS causes cannot be quantitatively studied.

REFERENCES

1. PANJER, H. H. "AIDS: Some Aspects of Modeling the Insurance Risk," *TSA* XLI (1989): 199–221.
2. RAMSAY, C. M. "AIDS and the Calculation of Life Insurance Functions," *TSA* XLI (1989): 393–422.

(AUTHOR'S REVIEW OF DISCUSSION)

COLIN M. RAMSAY:

I thank Dr. Seah and Dr. Chan for participating in the discussion of my paper.

Dr. Chan's model, which includes direct jumps to stages beyond the immediate next stage, is an interesting extension of my model; it is worth exploring further. One consequence of his model, however, is that lives will develop AIDS faster. Interestingly, according to this model, a life that is in stage 1 (that is, HIV +) can instantaneously "die of full-blown AIDS" in stage 5!

Dr. Seah's suggestions that age and duration should be included are well taken. I am currently working on extending my model to include both age and duration.