**Supplementary Online Material**

**Estimated HIV Incidence in the United States, 2003 – 2010**

Qiang Xia, MD, MPH, Armando Teixeira-Pinto, PhD, Lisa A. Forgione, MA,

Ellen W. Wiewel, DrPH, MHS, Sarah L. Braunstein, PhD, MPH, Lucia V. Torian, PhD

1. **Detailed steps for solving the equations**

$R=\frac{D\_{1}}{U\_{1}+I\_{1}}=\frac{D\_{2}}{U\_{2}+I\_{2}}=\frac{D\_{3}}{U\_{3}+I\_{3}}$ (Equation 1)

$U\_{2}=U\_{1}+I\_{1}-D\_{1}$ (Equation 2)

$U\_{3}=U\_{2}+I\_{2}-D\_{2}$ (Equation 3)

$I\_{2}=I\_{1}+ϵ\_{1}, where ϵ\_{1} is small$ (Equation 4)

$I\_{3}=I\_{2}+ϵ\_{2}, where ϵ\_{2} is small$ (Equation 5)

*R*: diagnosis rate, which is the probability of an HIV-infected individual being diagnosed in a year;[1](#_ENREF_1),[2](#_ENREF_2)

*Di*: number of new diagnoses in year *i*, where *i* = 1, 2, or 3;

*Ui*: number of undiagnosed persons at the beginning of year *i*, where *i* = 1, 2, or 3;

*Ii*: HIV incidence in year *i*, where *i* = 1, 2, or 3

1. $R=\frac{D\_{1}}{U\_{1}+I\_{1}}=\frac{D\_{2}}{U\_{2}+I\_{2}}=\frac{D\_{3}}{U\_{3}+I\_{3}}$
2. $\frac{D\_{1}}{U\_{1}+I\_{1}}=\frac{D\_{2}}{U\_{2}+I\_{2}}=\frac{D\_{3}}{U\_{3}+I\_{3}}$
3. $\frac{D\_{1}}{U\_{1}+I\_{1}}=\frac{D\_{2}}{(U\_{1}+I\_{1}-D\_{1})+I\_{2}}=\frac{D\_{3}}{(U\_{2}+I\_{2}-D\_{2})+I\_{3}}$
4. $\frac{D\_{1}}{U\_{1}+I\_{1}}=\frac{D\_{2}}{(U\_{1}+I\_{1}-D\_{1})+I\_{2}}=\frac{D\_{3}}{[(U\_{1}+I\_{1}-D\_{1})+I\_{2}-D\_{2}]+I\_{3}}$
5. $\frac{D\_{1}}{U\_{1}+I\_{1}}=\frac{D\_{2}}{U\_{1}+I\_{1}+I\_{2}-D\_{1}}=\frac{D\_{3}}{U\_{1}+I\_{1}+I\_{2}+I\_{3}-D\_{1}-D\_{2}}$
6. $\frac{D\_{1}}{U\_{1}+I\_{2}-ϵ\_{1}}=\frac{D\_{2}}{U\_{1}+(I\_{2}-ϵ\_{1})+I\_{2}-D\_{1}}=\frac{D\_{3}}{U\_{1}+(I\_{2}-ϵ\_{1})+I\_{2}+(I\_{2}+ϵ\_{2})-(D\_{1}+D\_{2})}$
7. $\frac{D\_{1}}{U\_{1}+I\_{2}-ϵ\_{1}}=\frac{D\_{2}}{U\_{1}+2I\_{2}-D\_{1}-ϵ\_{1}}=\frac{D\_{3}}{U\_{1}+3I\_{2}-(D\_{1}+D\_{2})-(ϵ\_{1}-ϵ\_{2})}$
8. $\frac{D\_{1}}{U\_{1}+I\_{2}}≈\frac{D\_{2}}{U\_{1}+2I\_{2}-D\_{1}}≈\frac{D\_{3}}{U\_{1}+3I\_{2}-(D\_{1}+D\_{2})}$

Solve for *U1* in equation $\frac{D\_{1}}{U\_{1}+I\_{2}}≈\frac{D\_{2}}{U\_{1}+2I\_{2}-D\_{1}}$

1. $\frac{D\_{1}}{U\_{1}+I\_{2}}≈\frac{D\_{2}}{U\_{1}+2I\_{2}-D\_{1}}$
2. $D\_{1}(U\_{1}+2I\_{2}-D\_{1})≈D\_{2}(U\_{1}+I\_{2})$
3. $D\_{1}U\_{1}+2D\_{1}I\_{2}-D\_{1}D\_{1}≈D\_{2}U\_{1}+D\_{2}I\_{2}$
4. $D\_{1}U\_{1}-D\_{2}U\_{1}≈D\_{2}I\_{2}-2D\_{1}I\_{2}+D\_{1}D\_{1}$
5. $(D\_{1}-D\_{2})U\_{1}≈(D\_{2}-2D\_{1})I\_{2}+D\_{1}D\_{1}$
6. $U\_{1}≈\frac{(D\_{2}-2D\_{1})I\_{2}+D\_{1}D\_{1}}{D\_{1}-D\_{2}}$

Substitute the value $\frac{(D\_{2}-2D\_{1})I\_{2}+D\_{1}D\_{1}}{D\_{1}-D\_{2}}$ for *U1* in equation $\frac{D\_{1}}{U\_{1}+I\_{2}}≈\frac{D\_{3}}{U\_{1}+3I\_{2}-(D\_{1}+D\_{2})}$

1. $\frac{D\_{1}}{\frac{(D\_{2}-2D\_{1})I\_{2}+D\_{1}D\_{1}}{D\_{1}-D\_{2}}+I\_{2}}≈\frac{D\_{3}}{\frac{(D\_{2}-2D\_{1})I\_{2}+D\_{1}D\_{1}}{D\_{1}-D\_{2}}+3I\_{2}-(D\_{1}+D\_{2})}$
2. $\frac{D\_{1}}{\left(D\_{2}-2D\_{1}\right)I\_{2}+D\_{1}D\_{1}+(D\_{1}-D\_{2})I\_{2}}≈\frac{D\_{3}}{\left(D\_{2}-2D\_{1}\right)I\_{2}+D\_{1}D\_{1}+3(D\_{1}-D\_{2})I\_{2}-(D\_{1}+D\_{2})(D\_{1}-D\_{2})}$
3. $\frac{D\_{1}}{-D\_{1}I\_{2}+D\_{1}D\_{1}}≈\frac{D\_{3}}{\left(D\_{1}-2D\_{2}\right)I\_{2}+D\_{2}D\_{2}}$
4. $\frac{1}{-I\_{2}+D\_{1}}≈\frac{D\_{3}}{\left(D\_{1}-2D\_{2}\right)I\_{2}+D\_{2}D\_{2}}$
5. $\left(D\_{1}-2D\_{2}\right)I\_{2}+D\_{2}D\_{2}≈-D\_{3}I\_{2}+D\_{1}D\_{3}$
6. $\left(D\_{1}-2D\_{2}+D\_{3}\right)I\_{2}≈D\_{1}D\_{3}-D\_{2}D\_{2}$
7. $I\_{2}≈\frac{D\_{1}D\_{3}-D\_{2}D\_{2}}{D\_{1}-2D\_{2}+D\_{3}}$
8. **An example – HIV incidence in the United States, 2003**

**Total**

Number of new diagnoses in 2002: D1 = 56715

Number of new diagnoses in 2003: D2 = 53106

Number of new diagnoses in 2004: D3 = 52758

HIV incidence in 2003: $I\_{2}≈\frac{D\_{1}D\_{3}-D\_{2}D\_{2}}{D\_{1}-2D\_{2}+D\_{3}}≈\frac{56715\*52758-53106\*53106}{56715-2\*53106+52758}≈\frac{171922734}{3261}≈52721$

**Male**

Number of new diagnoses in 2002: D1 = 41010

Number of new diagnoses in 2003: D2 = 38783

Number of new diagnoses in 2004: D3 = 38958

HIV incidence in 2003: $I\_{2}≈\frac{D\_{1}D\_{3}-D\_{2}D\_{2}}{D\_{1}-2D\_{2}+D\_{3}}≈\frac{41010\*38958-38783\*38783}{41010-2\*38783+38958}≈\frac{93546491}{2402}≈38945$

**Female**

Number of new diagnoses in 2002: D1 = 15705

Number of new diagnoses in 2003: D2 = 14323

Number of new diagnoses in 2004: D3 = 13800

HIV incidence in 2003: $I\_{2}≈\frac{D\_{1}D\_{3}-D\_{2}D\_{2}}{D\_{1}-2D\_{2}+D\_{3}}≈\frac{15705\*13800-14323\*14323}{15705-2\*14323+13800}≈\frac{11580671}{859}≈13482$

The sum of males and females (38,945 + 13,482 = 52,427) is 0.5577% smaller than the total ((52,427 – 52,721)/52,721 = -0.5577%). We require that the sum of estimated HIV incidence by sex equals the estimate of the total.[3](#_ENREF_3) In this case, the estimates for males and females need to be rescaled up by 0.5608% ((52,721 – 52,427)/52,427 = 0.5608%). The new estimate for males is 39,164 (38,945 \* (1 + 0.5608%) = 39,164), the new estimate for females is 13,557 (13,482 \* (1 + 0.5608%) = 13,557), and the sum (39,164 + 13,557 = 52,721) equals the total (52,721).

**References**

**1.** Xia Q, Kobrak P, Wiewel EW, Torian LV. The high proportion of late HIV diagnoses in the United States is likely to stay: findings from a mathematical model. *AIDS Care.* 2015;27(2):206-212.

**2.** Xia Q, Torian LV, Shepard CW. Limitations of indicators of HIV case finding. *Epidemiology.* 2015;26(1):e6-e8.

**3.** Murray CJ, Ortblad KF, Guinovart C, et al. Global, regional, and national incidence and mortality for HIV, tuberculosis, and malaria during 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet.* 2014;384(9947):1005-1070.