Gauging Regional Differences in the HIV Prevalence Rate Among Injection Drug Users in the U.S.

Jichuan Wang[*,1] and Brian Kelly[2]

[1]*Children's National Medical Center (CNMC), The George Washington University School of Medicine, USA*

[2]*Department of Sociology, Purdue University, USA*

Abstract: This article aims to introduce and demonstrate the application of the standardization and decomposition analysis (SDA) method to gauge differences in HIV prevalence rates among injection drug users (IDUs) across regions (Northeast, South, Midwest, and West) in the U.S. Using the SDA, the regional HIV prevalence rates were standardized and a rate difference between regions was decomposed into component effects, such as the "real" rate difference, and component effects attributed to differences in specific compositions of confounding factors. A total of 9,824 injection drug users (IDUs) retrieved from the national database of the National Institute on Drug Abuse's Cooperative Agreement for AIDS Community-Based Outreach/Intervention Research Program (COOP) projects constitute the sample for the study. A computer program DECOMP was used to implement the multi-population SDA.

Keywords: Standardization and decomposition analysis, DECOMP, HIV prevalence, injection drug users, disparity.

INTRODUCTION

Injection drug use remains a key pathway for HIV transmission and contributes to the diffusion of HIV globally. Since the onset of the HIV epidemic, scholars have estimated that drug injectors account for a significant proportion of the global burden of HIV [1]. Relative to the concentration of HIV within sexual networks in sub-Saharan Africa, the prevalence of HIV among drug injectors is higher in many regions of the world, such as the Middle East and East Asia, and is of particular concern in areas in which HIV prevalence is still rising, such as Eastern Europe and Central Asia [1]. While intervention efforts, notably forms of harm reduction such as needle exchange programs, have had a significant impact in reducing HIV transmission within drug injector networks, the risks faced by drug injectors remains considerable [2]. Precise estimates of the prevalence of HIV and other infectious diseases among drug injectors remain a critical public health objective. More precise estimates permit proper allocation of resources, facilitate the planning of harm reduction efforts, and direct health promotion and intervention initiatives as well as the provision of antiretroviral treatments [3]. In this study, we demonstrate application of standardization and decomposition analysis (SDA) to gauge regional differences in prevalence rates of HIV among injection drug users in the U.S., adjusting for confounding effects.

When comparing outcomes between populations or the same population at different time points, the effects of confounding factors should be taken into account in order to fully assess the circumstances. For example, it is possible for one population to have a crude death rate (the number of deaths occurring in a given year divided by the middle year population size) that is lower than another population's despite that the first population in fact has higher mortality (e.g., higher age-specific death rates). This paradox results from the fact that the first population has a larger proportion of its population in age groups (e.g., age 5-24) that are subject to lower mortality rates. That is, the difference in the observed crude death rates between the two populations is confounded by the differential age structures. Once age structure is standardized, the adjusted death rate of the first population would be certainly higher than that of the second population.

Standardization and decomposition analysis (SDA) is an analytical method used for outcome comparisons in demography and population studies [4-9]. For the above death rate example, the difference in the crude death rates between the two populations can be decomposed into two component effects: 1) the factor component effect attributed to different age structures; and 2) the rate effect attributed to the difference in mortality level (e.g., the age-specific death rates in this case). If another confounding factor (e.g., ethnicity) were taken into account, the difference in the crude death rate would be decomposed into three component effects: 1) factor component effect-1 effect due to the difference in age structure; 2) factor component effect-2 effect due to the difference in ethnic composition; and 3) the rate effect due to the difference in the factor specific (e.g., ethnicity-age specific) death rates.

SDA has some explicit advantages for data analysis. First, its results can be presented in a manner that is intuitively understandable. Outcome difference/change is decomposed into component effects that are attributed to "real" difference/change and effects of confounding factors; and the relative contributions of the component effects sum

*Address correspondence to this author at the George Washington University, Washington, DC, USA; Tel: 202-476-2978;
E-mails: jiwang@gwu.edu, jiwang@cnmc.org

1874-9410/14

up to 100%. These kinds of results are much easier than the statistical model parameter estimates (e.g., regression coefficients, odds ratio, hazard ratio) to understand, and in this regard the method is highly accessible to policy makers. Second, SDA is based on algebraic calculation, and thus it has no assumptions, such as multinormality, linear relationship, and observation independence that are often needed for statistical analyses. Third, SDA allows a wide range of outcome measures, such as rate, percentage, proportion, ratio, as well as arithmetic mean [2]. And four, SDA can be readily applied to longitudinal data with an unlimited number of time points.

Because SDA is based on algebraic calculation, traditionally, it does not provide significance testing for component effects. The lead author has developed a computer program, DECOMP, for SDA, which provides the opportunity to conduct significance testing for component effects using bootstrap methods [10]. The SDA and DECOMP were successfully applied to comparison of HIV seropositivity rates between male and female injection drug users (IDU) [10], and studying components of difference in HIV seropositivity rate among IDUs between low- and high-HIV-prevalence regions in the U.S. [11]. A multi-population SDA was conducted to examine regional differences in drug use practices (e.g., dichotomous and continuous measures of methamphetamine use) among rural stimulant users in Arkansas, Kentucky, and Ohio [12]. The results show that regional differences in observed measures of drug use were significantly confounded by socio-demographic factors. The improvement of the SDA method and development of the computer program have been academically accepted [13]. However, the applications of SDA to substance abuse and HIV prevention studies remain very limited. The present study aims to demonstrate the application of the multi-population SDA using real research data.

THE ALGEBRAIC EXPRESSION OF SDA

For a two-population comparison with only one confounding factor, the algebraic expression of SDA can be shown as follows [14, 15]:

$$R_{1.} = \sum_{j=1}^{J} \frac{N_{1j}R_{1j}}{N_{1.}} = \sum_{j=1}^{J} F_{1j}R_{1j} \quad (1)$$

$$R_{2.} = \sum_{j=1}^{J} \frac{N_{2j}R_{2j}}{N_{2.}} = \sum_{j=1}^{J} F_{2j}R_{2j} \quad (2)$$

where $R_{1.}$ denotes the observed rate (or mean if the outcome is a continuous measure) for Population 1; R_{1j} the observed factor-specific rate in the j^{th} category of the confounder with J categories(j=1, 2, … J) in Population 1; N_1 is the total number of cases in Population 1; N_{1j} specifies the number of cases in the j^{th} category of the confounder in Population 1; and $F_{1j}=N_{1j}/N_{1.}$ represents the proportion or relative frequency of the Population 1 members who fall into the j^{th} category of the confounder, and $\sum F_{1j}=1$. $R_{2.}$, R_{2j}, $N_{2.}$, N_{2j}, and F_{2j} are the equivalent notations for Population 2. In both Equations 1 and 2, the observed rate is expressed as a summation of weighted factor-specific rates: for instance, the weight is $F_{1j} =N_{1j}/N_{1.}$ for Population 1, and $F_{2j} =N_{2j}/N_{2.}$ for Population 2, which are the compositions of the confounder in each respective populations. The difference between the observed rates can be expressed accordingly:

$$R_{1.} - R_{2.} = \sum_{j=1}^{J} \frac{F_{1j}+F_{2j}}{2}\left(R_{1j} - R_{2j}\right) + \sum_{j=1}^{J} \frac{R_{1j}+R_{2j}}{2}\left(F_{1j} - F_{2j}\right) \quad (3)$$

Equation 3 shows that the difference between the two observed rates, ($R_{1.}$-$R_{2.}$), can be decomposed into two components: a rate effect (i.e., the first term in the equation) and a factor component effect (i.e., the second term in the equation). As shown in the first term on the right side of Equation 3, the composition of the confound or is standardized across populations; thus, the observed rate difference contained in this term can be considered having resulted from differential factor-specific rates between the populations under study. Therefore, we called it *rate effect*. In contrast, the second term on right side of Equation 3, where the factor-specific rate is standardized, represents the component in the crude rate difference that is attributed to differential factor compositions between the two populations. We call this term *factor component effect*, which describes the effect of the factor composition on the observed rate difference.

The traditional SDA could only deal with comparison of two populations with two confounding factors [1, 2]. The method was generalized by Das Gupta [14, 15] for multiple population comparisons with multiple confounding factors. In theory, the generalized SDA does not have a limit on the number of populations to compare and number of confounding factors to analyze. Note that when multiple populations (or a single population at multiple time-points) are analyzed, naive pair-wise comparisons between populations are inappropriate because the pair-wise comparison results usually lack internal consistency. In naive pair-wise comparisons, the estimate of a standardized rate for a specific population may not be consistent in different pair-wise comparisons. As a result, the difference in the standardized rates between population 1 and population 2 plus the difference between population 2 and population 3 may not equal the difference between population 1 and population 3, and so on. In other words, the component effects themselves may lack internal consistency [14, 15, 16]. To avoid such a problem in SDA with multiple populations or for the same population at multiple time-points, all pair-wise comparisons should be conducted simultaneously adjusting for internal inconsistence. The formulas for comparing populations 1 and 2 in the presence of populations 3, 4, …, and K is described as the following [14, 15]:

$$A_{1.23\ldots K} = \frac{\sum_{j=2}^{K} A_{1.j}}{(K-1)} + \frac{\sum_{i=2}^{K}\left(\sum_{j\neq 1,i}^{K} A_{i.j} - (K-2)A_{i.1}\right)}{K(K-1)} \quad (4)$$

$$A_{12.3\ldots K} = A_{12} - \frac{\sum_{i=3}^{K}\left(A_{12}+A_{2j}-A_{1j}\right)}{K} \quad (5)$$

where $A_{1.23\ldots K}$ and $A_{12.3\ldots K}$ are the standardized rate in population 1 and the effect of factor A, respectively, standardizing all other factors but A, when populations 1 and 2 are compared in presence of populations 3, 4, …, K. When populations 2 and 3 are compared in presence of population 1, 4, …, K, the corresponding standardized rate in population 2 and the effect of factor A would be $A_{2.31\ldots K}$ and $A_{23.14\ldots K}$, respectively. As a result, each population will have a

consistent standardized rate when standardization is conducted with respect to the same set of factors no matter which population it is compared with. It, therefore, solves the problem of internal inconsistency in component effect estimates in multiple population comparison. The standardized rates and component effects with respect to other factors can be calculated in the same way. That is, the same formulas apply to other factors regardless of how many factors are involved in the SDA [14, 15].

Table 1. HIV Prevalence Rate Among Injection Drug Users By Region and Site.

Region[1]	HIV Antibody Test	
	N[2]	HIV Prevalence Rate[3] (%)
Northeast		
New York	179	29.61
Philadelphia	554	13.00
Hartford	89	21.35
Subtotal	822	17.52
Midwest		
Detroit	684	10.82
Columbus/Dayton	1,036	1.45
St. Louis	365	1.64
Subtotal	2,085	4.56
South		
Houston	455	10.99
Miami	139	46.76
New Orleans	469	8.10
Lexington	328	3.96
Washington, DC	910	16.48
Durham/Wake Counties	254	12.99
San Antonio	284	3.87
Subtotal	2,839	12.68
West		
Oakland/Richmond	1,311	17.93
Anchorage	279	2.15
Flagstaff	86	1.16
Denver	223	3.59
Tucson	552	3.99
Portland	932	1.82
Long Beach	695	5.61
Subtotal	4,078	8.04
Total	9,824	9.44

Notes:
[1]US Census Regions.
[2]Number of IDUs who took voluntary and confidential HIV antibody tests at the baseline interview.
[3]The percentage of HIV positives among the IDUs in each sample was used as an estimate of HIV prevalence rate for that sample.

Although the mathematical formulas expressed in Equations 4 and 5 for multi-population comparisons are complicated, in computer program DECOMP multi-population SDA is implemented in the same way as two-population SDA, and its results are also interpreted in the same way [10, 12]. In applications of the current version of DECOMP, what one needs to do are: [1] open the program and input raw (text format) data; [2] specify the population variable (a categorical variable that has as many categories as the number of populations); [3] specify the outcome variable and select the confounding factors; and then [4] click the Run button.

The program results will show that each population will have a consistent standardized rate when standardization is conducted with respect to the same set of factors no matter which population it is compared with. It, therefore, solves the problem of internal inconsistency in component effect estimates in multi-population SDA. DECOMP allows an unlimited number of populations/samples for multiple comparisons; however, the number of confounding factors is limited up to 10.

SDA is traditionally applied using aggregated population data (e.g., contingency tables) and no significance testing is involved. In social science studies, survey data are often used and sampling variation must be taken into account. In order to make statistical inferences from survey data, significance tests of the component effects are needed in SDA. This challenge is completed by applying a bootstrapping technique in DECOMP [10]. As noted by Chevan & Sutherland: "*Wang et al. (2000) contributed to the enhancement of decomposition methods stemming from Das Gupta's work by developing tests of significance for decomposed rates using bootstrapping techniques to estimate standard errors*" [13, p.430].

DECOMP is a user-friendly Windows based computer program. Both grouped data (i.e., contingency tables) and individual data can be used for SDA in DECOMP. When individual data are analyzed, the outcome measure could be either a dichotomous or a continuous variable; when analyzing grouped data, the outcome measure could be a rate, proportion, percentage, ratio, or arithmetic mean. Please note, significance testing is only available with individual data because bootstrapping is conducted only the individual level. However, DECOMP has a utility function to convert an original contingency table into an individual data set if the outcome measure is a rate, percentage, proportion or ratio. The computer program is freely available to download online:www.wright.edu/~jichuan.wang/.

METHODS

Demonstration of SDA

In this study, we demonstrate how to apply SDA in real research using DECOMP. A total of 9,824 injection drug users (IDUs) retrieved from the large national database of *the National Institute on Drug Abuse's Cooperative Agreement for AIDS Community-Based Outreach/Intervention Research Program* (*COOP*) [17] were used for the demonstration. The outcome measure is dichotomous (i.e., 1-HIV test positive; 0-Otherwise); thus, the mean of the outcome in a regional population is an estimated HIV prevalence rate in the region population. The study standardizes and decomposes the

Table 2. Socio-Demographic Compositions and HIV Prevalence Rate by Region

Variable	Region												Total HIV[1]
	Northeast (n=822)			Midwest (n=2,085)			South (n=2,839)			West (n=4,078)			
	n	(%)	HIV[1]	n	(%)	HIV[1]	n	(%)	HIV[1]	n	(%)	HIV[1]	
Ethnicity													
Black	630	(76.64)	18.73	1,670	(80.10)	5.57	2337	(82.32)	14.72	1,923	(47.16)	14.09	12.59
White	192	(23.36)	13.54	415	(19.90)	0.48	502	(17.68)	3.19	2,155	(52.84)	2.65	3.09
(χ^2 P-value)	(0.0978)			(<0.0001)			(<0.0001)			(<0.0001)			(<0.0001)
Gender													
Female	174	(21.17)	16.67	508	(24.36)	6.10	798	(28.11)	15.16	1,182	(28.98)	8.12	10.41
Male	648	(78.83)	17.75	1,577	(75.64)	4.06	2,041	(71.89)	11.71	2,896	(71.02)	8.01	9.08
(χ^2 P-value)	(0.7393)					(0.0547)			(0.0129)			(0.9061)	(0.0450)
Age Group													
<30	77	(9.37)	15.58	98	(4.70)	1.02	253	(8.91)	3.95	469	(11.50)	4.05	4.68
30-39	337	(41.00)	20.77	749	(35.92)	5.47	1,102	(38.82)	11.89	1,590	(38.99)	7.74	9.66
40+	408	(49.64)	15.20	1,238	(59.38)	4.28	1,484	(52.27)	14.76	2,019	(49.51)	9.21	10.10
(χ^2 P-value)	(0.1230)			(0.1014)			(<0.0001)			(0.0009)			(<0.0001)
Education													
<High School	337	(41.00)	18.69	735	(35.25)	4.08	1086	(38.25)	15.19	1191	(29.21)	8.98	10.90
High School	337	(41.00)	18.40	808	(38.75)	4.70	1159	(40.82)	11.56	1758	(43.11)	8.59	9.48
College+	148	(18.00)	12.84	542	(26.00)	4.98	594	(20.92)	10.27	1129	(27.69)	6.20	7.34
(χ^2 P-value)	(0.2533)			(0.7239)			(0.0049)			(0.0257)			(<0.0001)

Notes:

[1]Difference in HIV prevalence rate among demographic groups was tested using χ^2 statistics.

differences in the HIV prevalence rates between four U.S. geographic regions (Northeast, Midwest, South, and West). For the purpose of simplicity, only a limited number of socio-demographic factors, such as ethnicity, age, gender and education, are considered as confounding factors in the study. The sample consists of 66.8% African Americans and 33.2% Whites; the majority of the sample were male (72.9%); and mean age of the sample was 39.8.

RESULTS

The descriptive statistics for the sample and the estimates of HIV prevalence rates among IDUs are shown in Table **1** by region and research site. The HIV prevalence rate was high in the Northeast (17.52%) and the South (12.68%), moderate in the West (8.04%), and low in the Midwest (4.56%). Overall, socio-demographic characteristics, such as ethnicity, gender, age, and education, were significantly associated with the HIV prevalence rate (see the last column of Table **2**). The HIV prevalence rate was much higher among Black IDUs than among White IDUs across the regions (though only marginally significant in the Northeast). The gender effect was only significant in the South, while age and education had significant effects in the South and West. Notably, the compositions of the socio-demographic factors, ethnicity in particular, vary substantially across regions. For example, only 47.16% of the IDUs in the West were Blacks, while the corresponding figures were 76.64% in the Northeast, 80.10% in the Midwest, and 82.32% in the South, respectively.

The results of SDA are shown in Table **3**. In the first panel of the table, we show results comparing HIV prevalence rates between the Northeast and Midwest regions. The observed HIV prevalence rate was about 12.91% higher in the Northeast than in the Midwest. Significance testing for the component effects was conducted using the T-test, where the standard error of the difference was estimated based on 1000 bootstrap resamples. Adjusting for the confounding factors, the regional rate difference (12.81%) remains almost unchanged; this indicates that the socio-demographic factors do not confound the difference of HIV prevalence rates between the Northeast and the Midwest regions. Consistently, the SDA results show that all of the factor component effects are very small and not statistically significant.

The HIV prevalence rates between the Northeast and the West decreased from 9.47% to 6.51% after adjusting for the

Table 3. Results of Multi-Population Standardization and Decomposition Analysis based on 1000 Bootstrap Resamples

	Standardization		Decomposition		
	Northeast	Midwest	Difference		Percent Distribution of Effect (%)
			Diff	S.E.	
Ethnicity	0.1112	0.1131	-0.0019	0.0015	-1.4722
Gender	0.1069	0.1067	0.0002	0.0009	0.1550
Age	0.1078	0.1077	0.0001	0.0011	0.0775
Education	0.1090	0.1063	0.0027	0.0011	2.0920
Adjusted rate	0.1697	0.0416	0.1281	0.0094	99.2549
Observed rate	0.1752	0.0462	0.1291	0.0096	100.0297
	Northeast	South	Difference		Percent Distribution of Effect (%)
			Diff	S.E.	
Ethnicity	0.1112	0.1185	-0.0073	0.0017	-15.0764
Gender	0.1069	0.1074	-0.0005	0.0011	-1.0326
Age	0.1078	0.1066	0.0012	0.0009	2.4783
Education	0.1090	0.1075	0.0015	0.0010	3.0979
Adjusted rate	0.1697	0.1162	0.0535	0.0102	110.4911
Observed rate	0.1752	0.1268	0.0484	0.0103	99.9583
	Northeast	West	Difference		Percent Distribution of Effect (%)
			Diff	S.E.	
Ethnicity	0.1112	0.0855	0.0257	0.0028	27.1254
Gender	0.1069	0.1076	-0.0007	0.0011	-0.7388
Age	0.1078	0.1063	0.0015	0.0009	1.5832
Education	0.1090	0.1057	0.0033	0.0012	3.4830
Adjusted rate	0.1697	0.1046	0.0651	0.0092	68.7107
Observed rate	0.1752	0.0805	0.0947	0.0094	99.9524
	Midwest	South	Difference		Percent Distribution of Effect (%)
			Diff	S.E.	
Ethnicity	0.1131	0.1185	-0.0054	0.0012	6.6963
Gender	0.1067	0.1074	-0.0007	0.0006	0.8680
Age	0.1077	0.1066	0.0011	0.0007	-1.3641
Education	0.1063	0.1075	-0.0012	0.0007	1.4881
Adjusted rate	0.0416	0.1162	-0.0746	0.0063	92.5082
Observed rate	0.0462	0.1268	-0.0806	0.0064	99.9486
	Midwest	West	Difference		Percent Distribution of Effect (%)
			Diff	S.E.	
Ethnicity	0.1131	0.0855	0.0276	0.0020	-80.4277
Gender	0.1067	0.1076	-0.0009	0.0006	2.6226
Age	0.1077	0.1063	0.0014	0.0008	-4.0797
Education	0.1063	0.1057	0.0006	0.0009	-1.7484
Adjusted rate	0.0416	0.1046	-0.0630	0.0059	183.5849
Observed rate	0.0462	0.0805	-0.0343	0.0057	99.9518

(Table 3) contd…..

	South	West	Difference		Percent Distribution of Effect (%)
			Diff	S.E.	
Ethnicity	0.1185	0.0855	0.0330	0.0021	71.2359
Gender	0.1074	0.1076	-0.0002	0.0007	-0.4317
Age	0.1066	0.1063	0.0003	0.0006	0.6476
Education	0.1075	0.1057	0.0018	0.0008	3.8856
Adjusted rate	0.1162	0.1046	0.0116	0.0073	25.0405
Observed rate	0.1268	0.0805	0.0463	0.0071	99.9462

confounding factors (see the third panel of Table **3**). That is, assuming the same socio-demographic compositions, the regional difference in HIV prevalence rate would be about 31.26% smaller. The adjusted rate difference reflects the factor-specific rate difference, which accounts for about 68.71% of the observed regional difference in HIV prevalence rate (see the last column of the panel in Table **3**). Ethnic composition had a significant confounding effect (t-ratio=0.0257/0.0028=9.18), accounting for about 27.13% of the observed prevalence rate difference. The contributions of the factors of age and education to the crude rate difference were very limited(1.58% and 3.48%, respectively) and statistically insignificant. The confounding effect attributed to gender composition is -0.74%. This indicates that the observed rate difference would be slightly smaller if gender composition were not controlled. However, this confounding effect is not statistically significant (t-ratio=-0.0007/0.0011= 0.63).

Ethnicity's confounding effect is statistically significant in five of the six pair-wise regional comparisons (see Table **3**); however, the confounding effect was positive sometimes (e.g., Northeast *vs* West; Midwest *vs* West; and South *vs* West) and negative sometimes (e.g., Northeast *vs* Midwest; Northeast *vs* South; and Midwest *vs* South). A positive confounding effect means that the rate difference between regions would be enlarged if the confounding effect were not controlled; on the contrary, a negative confounding effect indicates the extent to which the rate difference would be narrowed if the confounding effect were not controlled.

Education shows significant confounding effects on differences in HIV prevalence rates in the following regional comparisons: Northeast *vs* Midwest (t-ratio=0.0027/0.0011= 2.45); Northeast *vs* West (t-ratio=0.0033/0.0012=2.75); and South *vs* West (t-ratio=0.0018/0.0008=2.25). Gender and age had no significant confounding effects on any regional differences in HIV prevalence rates because age structure and gender compositions do not vary much across regions (see Table **2**).

DISCUSSION

As injection drug use remains an important component of the global HIV epidemic, the precise estimation of HIV prevalence across regions remains critical [1]. The results of this study show that ethnicity and education are important confounding factors in HIV prevalence rate comparisons among injection drug users across different U.S. regions. It is important to remember that decomposition of outcome differences using SDA is not equivalent to analyzing variance of a dependent variable in a regression model. A variable may significantly explain the variation of a dependent variable in regression, but may not have a significant confounding effect in SDA. For example both binary and multivariate statistics may show a significant relationship between ethnicity and an outcome of interest; however, ethnicity would have no significant confounding effect on the outcome difference between populations if ethnic composition does not vary much across the populations under study.

In the present study, ethnic composition was substantially different in the Northeast, South, and West regions of the U.S.; as such, the confounding effect of ethnicity was statistically significant when comparing these regions. Given that the ethnic composition between the Northeast and Midwest regions was not very different, the ethnic component effect in comparison of crude HIV prevalence rate between the two regions is not statistically significant (t-ratio=0.0019/ 0.0015=1.26). The consideration of these differences may enable more precise targeting of the at-risk population. Certainly, these considerations are important with the provision of harm reduction services and intervention efforts that are tailored to meet the needs of the population.

The SDA method is applicable to comparing different types of outcome measures, such as rate, percentage, proportion, ratio, and arithmetic mean among multiple populations/samples. It has wide applications within the study of both substance use and HIV/AIDS, as outcome comparisons or disparities are significant concerns within these areas. The SDA method can also be readily used to analyze outcome change and confounding effects on the change in longitudinal studies. In short, SDA is a useful analytical method for comparing outcomes in substance abuse and HIV studies. It provides an opportunity of viewing and interpreting outcome differences among different populations from a different perspective.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

Declared none.

REFERENCES

[1] UNAIDS. Global report: UNAIDS report on the global AIDS epidemic 2012. New York: United Nations 2012.

[2] Strathdee SA, Vlahov D. The effectiveness of needle exchange programs: A review of the science. AID Sci 2001; 1: 1-33

[3] Aceijas C, Stimson GV, Hickman M, Rhodes T. Global overview of injecting drug use and HIV infection among injecting drug users. AIDS 2004; 18: 2295-303.

[4] Kitagawa EM. Components of a difference between two rates. J Am Stat Assoc 1955; 50: 1168-94.

[5] Kitagawa EM. Standardized comparisons in population research. Demography 1964; 1: 296-315.

[6] Pullum TW. Standardization (world fertility survey technical bulletins, No. 597), Voorburg, Netherlands: International Statistical Institutes 1978.

[7] United Nations. The Methodology of measuring the impact of family planning programs, manual IX, population studies, No. 66. New York: United Nations 1979.

[8] Zheng H, Landb KC. Composition and decomposition in US gender-specific self-reported health disparities, 1984-2007. Soc Sci Res 2012; 41: 477-88.

[9] Trovato F, Odynak D. Sex differences in life expectancy incanada: immigrant and native-bornpopulations. J Biosoc Sci 2011; 43: 353-68.

[10] Wang J, Rahman A, Siegal HA, Fisher JH. Standardization and decomposition of rates: Useful analytic techniques for behavior and health studies. Behav Res Methods Instr Comput 2000; 32: 357-66.

[11] Wang J. Components of difference in HIV seropositivity rate among injection drug users between low and high HIV prevalence regions. AIDS Behav 2003; 7: 1-8.

[12] Wang J, Carlson RG, Falck RS, Leukefeld C, Booth BM. Multiple sample standardization and decomposition analysis: An application to comparisons of methamphetamine use among rural drug users in three American states. Stat Med 2007; 26: 3612-23.

[13] Chevan A, Sutherland M. Revisiting Das Gupta: refinement and extension of standardization and decomposition. Demography 2009; 46: 429-49.

[14] Das GP. Decomposition of the difference between two rates and its consistency when more than two populations are involved. Math Populat Stud 1991; 3: 105-25.

[15] Das GP. Standardization and decomposition of rates: A user's manual. U.S. Bureau of the Census, current population reports, series p23-186. Washington, D.C.: U.S. Government Printing Office 1993.

[16] Wang J. Some issues on application of standardization and decomposition analysis. Open Demogr J 2008; 1: 15-7.

[17] Needle R, Fisher DG, Weatherby N, *et al*. Reliability of self-reported HIV risk behaviorsof drug users. Psychol Addict Behav 1995; 9: 242-50.

Received: July 8, 2013 Revised: December 18, 2013 Accepted: January 2, 2014