DENVER HIV RISK SCORE

The Denver HIV Risk Score (DHRS) is an empirically-derived clinical prediction instrument originally developed in 2012 using approximately 100,000 observations, over 500 newly-diagnosed HIV infected individuals, and approximately 50 candidate predictor variables from a public health department sexually transmitted diseases clinic in Denver, Colorado. After extensive derivation and modeling,[1] the original DHRS included the following 8 variables: patient demographics (age, gender, and race/ethnicity), sexual behaviors (sex with males, vaginal intercourse, and receptive anal intercourse), injection drug use, and past HIV test. During initial testing, this original DHRS performed well in both internal validation and limited external validation (using emergency department populations from the University of Cincinnati Medical Center, Cincinnati, Ohio and Johns Hopkins Hospital, Baltimore, Maryland).[1,2] Additionally, the DHRS was recently used to guide targeted HIV screening in an emergency department (ED) and when compared to nontargeted screening was found to be more strongly associated with identification of newly-diagnosed HIV-infected patients.[3] However, given the potential sensitivity of some of the questions, the original DHRS was refined to exclude "vaginal intercourse" and "receptive anal intercourse" without modification to the scores of the remaining variables (**Supplemental Table 1**). As such, the refined DHRS (hereafter referred to as simply "the DHRS") includes 6 variables, 3 demographic variables and 3 risk characteristics, each with distinct assigned points based on regression coefficients from the original multivariable modeling.[1] It is important to note that age categories are mutually exclusive and that gender is independent of the "sex with a male" variable. For instance, males who have sex with males

score +43 points without consideration of the other risk score variables; alternatively, females who have sex with males score +22 points. The DHRS ranges from -4 to +73, with scores ≥30 representing patients at increased risk for HIV infection, and scores ≥50 representing the highest risk for HIV infection.

ADDITIONAL ANALYSES

In addition to the primary analyses reported in the main article, several additional secondary analyses were performed.

First, analyses were performed stratifying by geographic region as defined by the United States Census Bureau.[4] The overall prevalence of newly-diagnosed HIV infection among the five geographic regions of the United States ranged from 0.53% in the Northeast to 0.90% in Puerto Rico and the United States Virgin Islands, and HIV prevalence within each DHRS stratum was similar across geographic regions (**Supplemental Tables 2 and 3; Supplemental Figure 1**).

Second, analyses were performed to assess secular trend across the three-year study period, while also stratifying by geographic region (**Supplemental Table 4**). Although the prevalence of newly-diagnosed HIV infection decreased slightly from 2008 through 2010, this was driven most by testing performed in the South. There was variability in the prevalence of newly-diagnosed HIV infection when stratified by year and geographic regions, and given the similar performance of the DHRS across geographic regions, it is unlikely any secular trends identified in this dataset would impact the performance of the score.

Third, analyses were performed to assess if regional heterogeneity would change the relative weights of DHRS variables, therefore varying the performance of the DHRS according

to region. The cohort was subset according to region and all DHRS variables were included in new, separate multivariable logistic regression models to determine if regression coefficients changed significantly between regions. A risk score for each region was then created by multiplying each model's regression coefficients by 10 and rounding them to the nearest integer. Risk groups were then categorized according to the same cut points used when developing the original DHRS. We identified modest differences with respect to the weights of the DHRS variables (**Supplemental Tables 5 and 6**). The largest difference between regions was injection drug use (i.e., +10 for the Northeast region and +1 for the Midwest region), whereas age, gender, race/ethnicity, sex with a male, and past HIV test remained relatively stable across regions. Given these results, the performance of the DHRS within each region appeared to perform similarly to its performance when aggregated across all regions (**Supplemental Figures 2 – 6**).

Fourth, best-case and worst-case sensitivity analyses were performed to estimate the effect of missing data on complete-case results. Other methods of handling missing (e.g., multiple imputation) were not feasible given the limited number of variables in the dataset.[5,6] During the three-year study period, 6,768,987 total testing events were reported to the CDC; of these, 43,234 (0.6%) were newly-identified with HIV infection. Of the 6,768,987 observations, 4,830,941 (71%) were reported with complete data, and represent the principal cohort for the analyses described in the main article. **Supplemental Table 7** shows the distribution of the DHRS variables and proportions of missing data among the full cohort, stratified by outcomes. The largest proportion of missing data existed for injection drug use and past HIV test, although in general, less missingness existed for patients who were newly-diagnosed with HIV infection. **Supplemental Table 8** shows complete case results and sensitivity analyses where missing data

were coded using best-case and worst-case values. **Supplemental Figure 7** shows the possible range of the performance of the DHRS for the sensitivity analyses.

Fourth, calibration was assessed by comparing predicted versus observed HIV prevalences, and discrimination by estimating the area under the receiver operating characteristics (ROC) curve. The DHRS demonstrated excellent calibration (regression slope: 1.09 for the primary outcome, 1.10 for the secondary outcome; and an R^2 of 0.99 for the primary outcome, 0.98 for the secondary outcome) (**Supplemental Figure 8**) and discrimination (area under the ROC curve of 0.77 [95% CI: 0.77 – 0.77] for the primary outcome, and 0.74 [95% CI: 0.73 – 0.74]) for the secondary outcome (**Supplemental Figure 9**).

REFERENCES

1. Haukoos JS, Lyons MS, Lindsell CJ, et al. Derivation and validation of the Denver Human Immunodeficiency Virus (HIV) risk score for targeted HIV screening. Am J Epidemiol 2012;175:838-46.
2. Hsieh YH, Haukoos JS, Rothman RE. Validation of an abbreviated version of the Denver HIV Risk Score for prediction of HIV infection in an urban ED. The American Journal of Emergency Medicine 2014;32:775-9.
3. Haukoos JS, Hopkins E, Bender B, Sasson C, Al-Tayyib AA, Thrun MW. Comparison of enhanced targeted rapid HIV screening using the Denver HIV risk score to nontargeted rapid HIV screening in the emergency department. Ann Emerg Med 2013;61:353-61.
4. Census Regions of the United States. (Accessed July 31, 2013, at http://apps.bts.gov/publications/america_on_the_go/us_business_travel/html/figure_02.html.)
5. Newgard CD, Haukoos JS. Advanced statistics: missing data in clinical research--part 2: multiple imputation. Acad Emerg Med 2007;14:669-78.
6. Haukoos JS, Newgard CD. Advanced statistics: missing data in clinical research--part 1: an introduction and conceptual framework. Acad Emerg Med 2007;14:662-8.

SUPPLEMENTAL Table 1. The Denver HIV Risk Score.

Variable	Original Score	Refined Score
Age		
22-25 or 55-60 years	+4	+4
26-32 or 47-54 years	+10	+10
33-46 years	+12	+12
Gender		
Male	+21	+21
Race/Ethnicity		
Black	+9	+9
Hispanic	+3	+3
Sexual Practices		
Sex with a male	+22	+22
Vaginal intercourse	-10	-
Receptive anal intercourse	+8	-
Other Risks		
Injection drug use	+9	+9
Past HIV test	-4	-4

*Reference groups that score zero are: Age, <22 or >60 years; Gender, female; Race/Ethnicity, white or "other" (defined as American or Alaskan Native, Native Hawaiian, or non-Hawaiian Pacific Islander). Dashes represent exclusion of the variable from the refined score. The refined score ranges from -4 to +73 with risk groups stratified as <20 (very low risk), 20 – 29 (low risk), 30 – 39 (moderate risk), 40 – 49 (high risk, and ≥50 (very high risk).

SUPPLEMENTAL Table 2. Prevalence of HIV infection stratified by geographic region (N=4,830,941), CDC PEMS data, 2008 – 2010.

Geographic Region	Newly-Diagnosed HIV Infection		Confirmed HIV Positive Test Result	
	n	(%)	n	(%)
Northeast (N=1,219,070)	6,452	(0.53)	8,322	(0.68)
South (N=2,654,370)	16,561	(0.62)	26,936	(1.01)
Midwest (N=379,345)	2,265	(0.60)	2,677	(0.71)
West (N=492,807)	4,033	(0.82)	5,204	(1.06)
Puerto Rico & USVI (N=85,349)	769	(0.90)	1,374	(1.61)

Abbreviations: CDC = Centers for Disease Control and Prevention; HIV = human immunodeficiency virus; PEMS = Program Evaluation and Monitoring System; USVI = United States Virgin Islands.
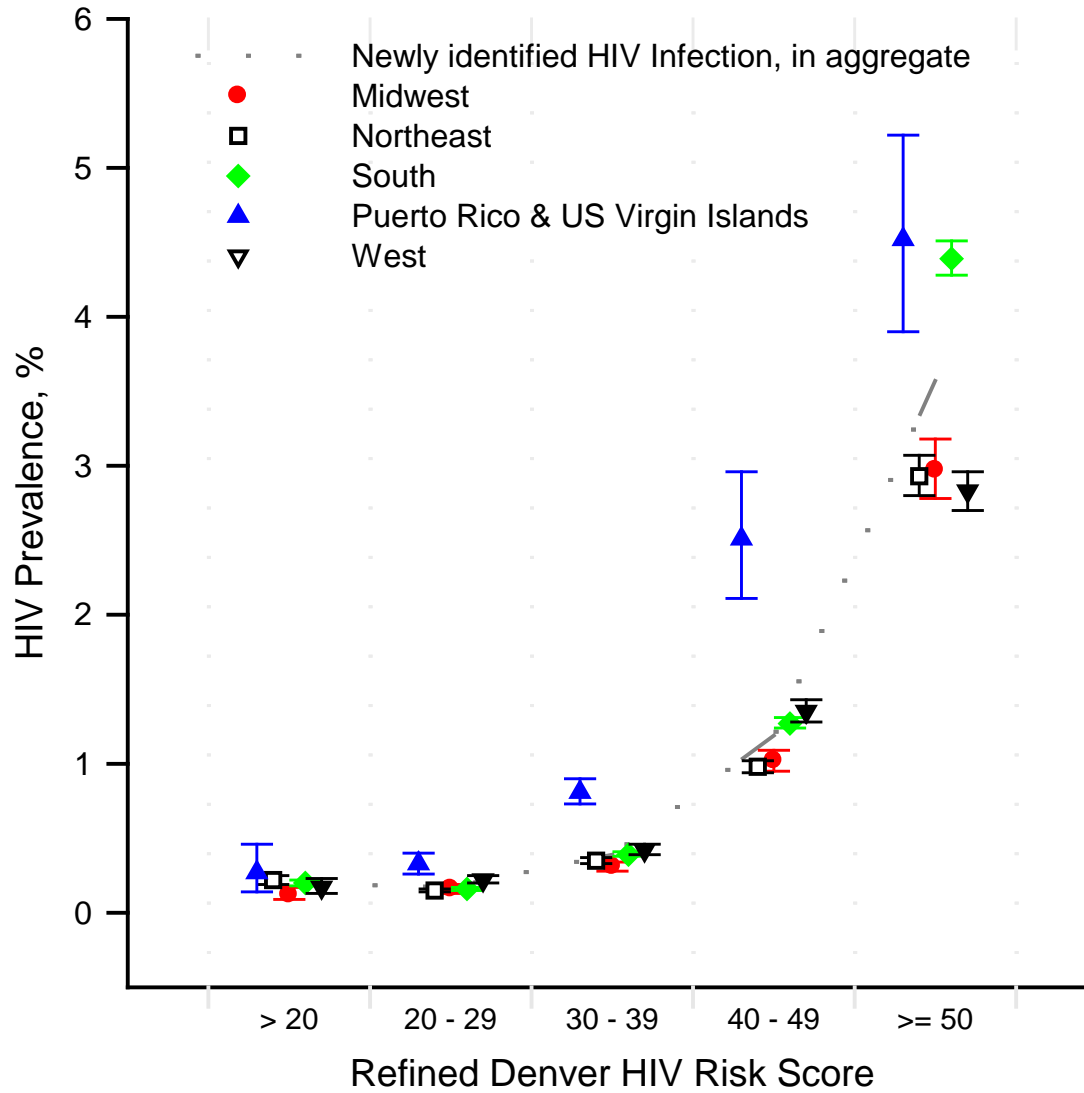
SUPPLEMENTAL Table 3. Performance of the Denver HIV Risk Score, stratified by geographic region (N=4,830,941), CDC PEMS data, 2008 - 2010.

DHRS	Newly-Diagnosed HIV Infection		Confirmed HIV Positive Test Result	
	Prevalence, %	(95 % CI)	Prevalence, %	(95 % CI)
All Regions				
< 20	0.20	(0.19 – 0.20)	0.50	(0.48 – 0.52)
20 – 29	0.17	(0.16 – 0.17)	0.28	(0.27 – 0.29)
30 – 39	0.39	(0.38 – 0.40)	0.62	(0.61 – 0.63)
40 – 49	1.19	(1.16 – 1.21)	1.64	(1.61 – 1.66)
≥ 50	3.57	(3.50 – 3.65)	4.79	(4.71 – 4.87)
Northeast Region Only				
< 20	0.22	(0.19 – 0.25)	0.31	(0.28 – 0.35)
20 – 29	0.15	(0.13 – 0.16)	0.21	(0.19 – 0.22)
30 – 39	0.35	(0.33 – 0.37)	0.46	(0.44 – 0.48)
40 – 49	0.98	(0.94 – 1.02)	1.23	(1.18 – 1.27)
≥ 50	2.93	(2.79 – 3.07)	3.69	(3.53 – 3.84)
South Region Only				
< 20	0.20	(0.18 – 0.22)	0.63	(0.60 – 0.66)
20 – 29	0.16	(0.15 – 0.17)	0.30	(0.29 – 0.32)
30 – 39	0.39	(0.38 – 0.41)	0.69	(0.68 – 0.71)
40 – 49	1.27	(1.24 – 1.31)	1.90	(1.85 – 1.94)
≥ 50	4.39	(4.28 – 4.51)	6.31	(6.17 – 6.45)
Midwest Region Only				
< 20	0.12	(0.09 – 0.17)	0.19	(0.14 – 0.24)
20 – 29	0.16	(0.13 – 0.19)	0.20	(0.18 – 0.23)
30 – 39	0.31	(0.28 – 0.34)	0.39	(0.36 – 0.42)
40 – 49	1.02	(0.95 – 1.09)	1.18	(1.10 – 1.26)
≥ 50	2.97	(2.78 – 3.18)	3.39	(3.18 – 3.60)
West Region Only				
< 20	0.17	(0.13 – 0.23)	0.34	(0.28 – 0.41)
20 – 29	0.22	(0.20 – 0.25)	0.32	(0.29 – 0.35)
30 – 39	0.42	(0.39 – 0.46)	0.60	(0.56 – 0.64)
40 – 49	1.35	(1.28 – 1.43)	1.70	(1.61 – 1.78)

≥ 50	2.83	(2.70 – 2.96)	3.44	(3.30 – 3.59)
Puerto Rico & USVI Only				
< 20	0.27	(0.14 – 0.46)	0.75	(0.53 – 1.04)
20 – 29	0.33	(0.26 – 0.40)	0.60	(0.52 – 0.70)
30 – 39	0.81	(0.73 – 0.90)	1.54	(1.43 – 1.67)
40 – 49	2.51	(2.11 – 2.96)	4.35	(3.82 – 4.92)
≥ 50	4.52	(3.90 – 5.22)	6.83	(6.06 – 7.66)

Abbreviations: DHRS = Denver HIV Risk Score; CDC = Centers for Disease Control and Prevention; HIV = human immunodeficiency virus; PEMS = Program Evaluation and Monitoring System; USVI = United States Virgin Islands.

SUPPLEMENTAL Figure 1. Prevalence of newly identified human immundeficiency virus (HIV) infection within each risk score category in the validation sample, stratified by geographical region, CDC PEMS data, 2008 - 2010. The refined Denver HIV Risk Score ranges from -4 to +73. Bars, 95% confidence interval.

SUPPLEMENTAL Table 4. Prevalence of newly-diagnosed HIV infection stratified by year and geographic region (N=4,830,941), CDC PEMS data, 2008 – 2010.

Geographic Region	2008 (N=1,477,672)		2009 (N=1,682,982)		2010 (N=1,670,287)	
	n	(%)	n	(%)	n	(%)
All Regions	10,251	(0.69)	10,090	(0.60)	9,739	(0.58)
Northeast Region Only	1,810	(0.53)	2,045	(0.46)	2,597	(0.60)
South Region Only	5,835	(0.74)	5,590	(0.61)	5,136	(0.54)
Midwest Region Only	909	(0.61)	765	(0.64)	591	(0.54)
West Region Only	1,498	(0.84)	1,461	(0.85)	1,074	(0.75)
Puerto Rico & USVI Only	199	(0.80)	229	(0.82)	341	(1.05)

Abbreviations: CDC = Centers for Disease Control and Prevention; HIV = human immunodeficiency virus; PEMS = Program Evaluation and Monitoring System; USVI = United States Virgin Islands.

SUPPLEMENTAL Table 5. Weights of refined DHRS variables according to geographic region (N=4,830,941), CDC PEMS data, 2008-2010.

Variable	DHRS	Northeast Region Only	South Region Only	Midwest Region Only	West Region Only	PR & USVI* Only
Age						
< 22 or > 60 years	ref	ref	ref	ref	ref	ref
22-25 or 55-60 years	+4	+6	+5	+1	+5	+5
26-32 or 47-54 years	+10	+7	+8	+2	+6	+10
33-46 years	+12	+10	+10	+4	+7	+12
Gender						
Female	ref	ref	ref	ref	ref	ref
Male	+21	+21	+25	+26	+22	+20
Race/Ethnicity						
Black	+9	+11	+10	+10	+7	-
Hispanic	+3	+5	+3	+1	+5	-
Other† or White	ref	ref	ref	ref	ref	-
Sexual Practices						
Sex with a male	+22	+19	+21	+22	+20	+15
Other Risks						
Injection drug use	+9	+10	+5	+1	+5	+8
Past HIV test	-4	-4	-3	0	-2	-3

Abbreviations: DHRS = Denver HIV Risk Score; CDC = Centers for Disease Control and Prevention; PEMS = Program Evaluation and Monitoring System.

* Model does not include race/ethnicity due to near complete separation of the values of these variables.

†Represents American or Alaskan Native, Native Hawaiian, or non-Hawaiian Pacific Islander.

SUPPLEMENTAL Table 6. Performance of remodeled DHRS by geographic region (N=4,830,941), CDC PEMS data, 2008-2010.

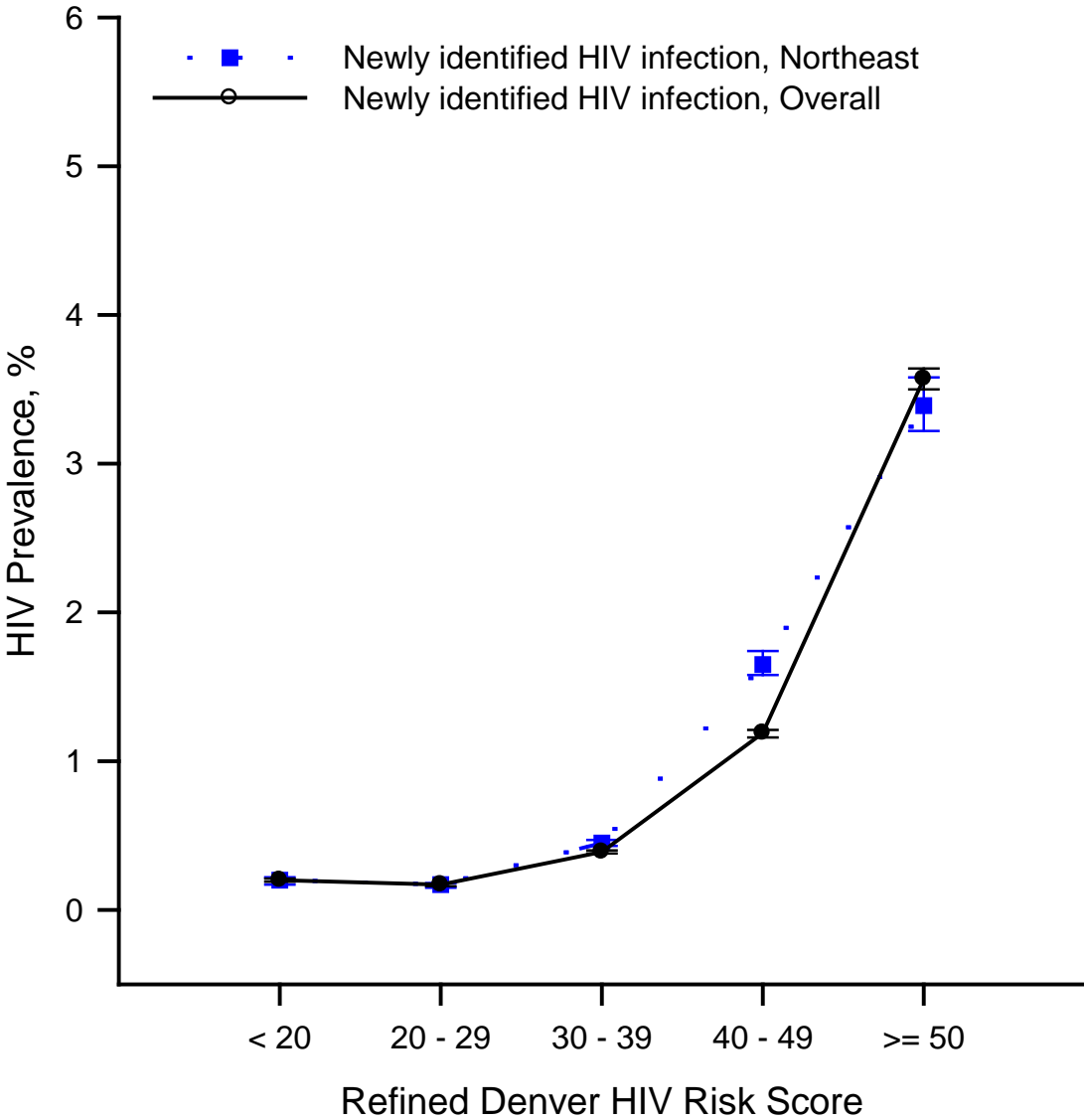| DHRS	New HIV Infections		Confirmed HIV Test Results	
	Prevalence, %	(95 % CI)	Prevalence, %	(95 % CI)
Northeast Region Only				
< 20	0.20	0.17 – 0.22	0.27	0.25 – 0.30
20 – 29	0.17	0.15 – 0.18	0.25	0.23 – 0.26
30 – 39	0.45	0.43 – 0.47	0.59	0.57 – 0.61
40 – 49	1.66	1.58 – 1.74	2.04	1.95 – 2.13
≥ 50	3.39	3.22 – 3.58	4.21	4.01 – 4.41
South Region Only				
< 20	0.21	0.19 – 0.23	0.65	0.62 – 0.69
20 – 29	0.13	0.13 – 0.14	0.25	0.24 – 0.27
30 – 39	0.39	0.37 – 0.40	0.68	0.67 – 0.70
40 – 49	0.97	0.94 – 1.00	1.52	1.48 – 1.55
≥ 50	4.13	4.04 – 4.23	5.86	5.74 – 5.97
Midwest Region Only				
< 20	0.16	0.11 – 0.22	0.21	0.16 – 0.28
20 – 29	0.15	0.13 – 0.18	0.19	0.16 – 0.21
30 – 39	0.35	0.32 – 0.38	0.44	0.41 – 0.48
40 – 49	0.94	0.84 – 1.04	1.09	0.98 – 1.20
≥ 50	2.83	2.68 – 3.00	3.23	3.06 – 3.40
West Region Only				
< 20	0.18	0.13 – 0.24	0.34	0.28 – 0.42
20 – 29	0.24	0.22 – 0.27	0.36	0.33 – 0.39
30 – 39	0.48	0.44 – 0.52	0.66	0.62 – 0.71
40 – 49	1.80	1.71 – 1.89	2.26	2.16 – 2.36
≥ 50	3.47	3.28 – 3.66	4.11	3.91 – 4.31
Puerto Rico & USVI Only				
< 20	0.28	0.11 – 0.19	0.64	0.53 – 0.75
20 – 29	0.59	0.53 – 0.67	1.18	1.09 – 1.29
30 – 39	1.91	1.67 – 2.17	3.16	2.86 – 3.49
40 – 49	4.68	4.04 – 5.39	6.68	5.92 – 7.51

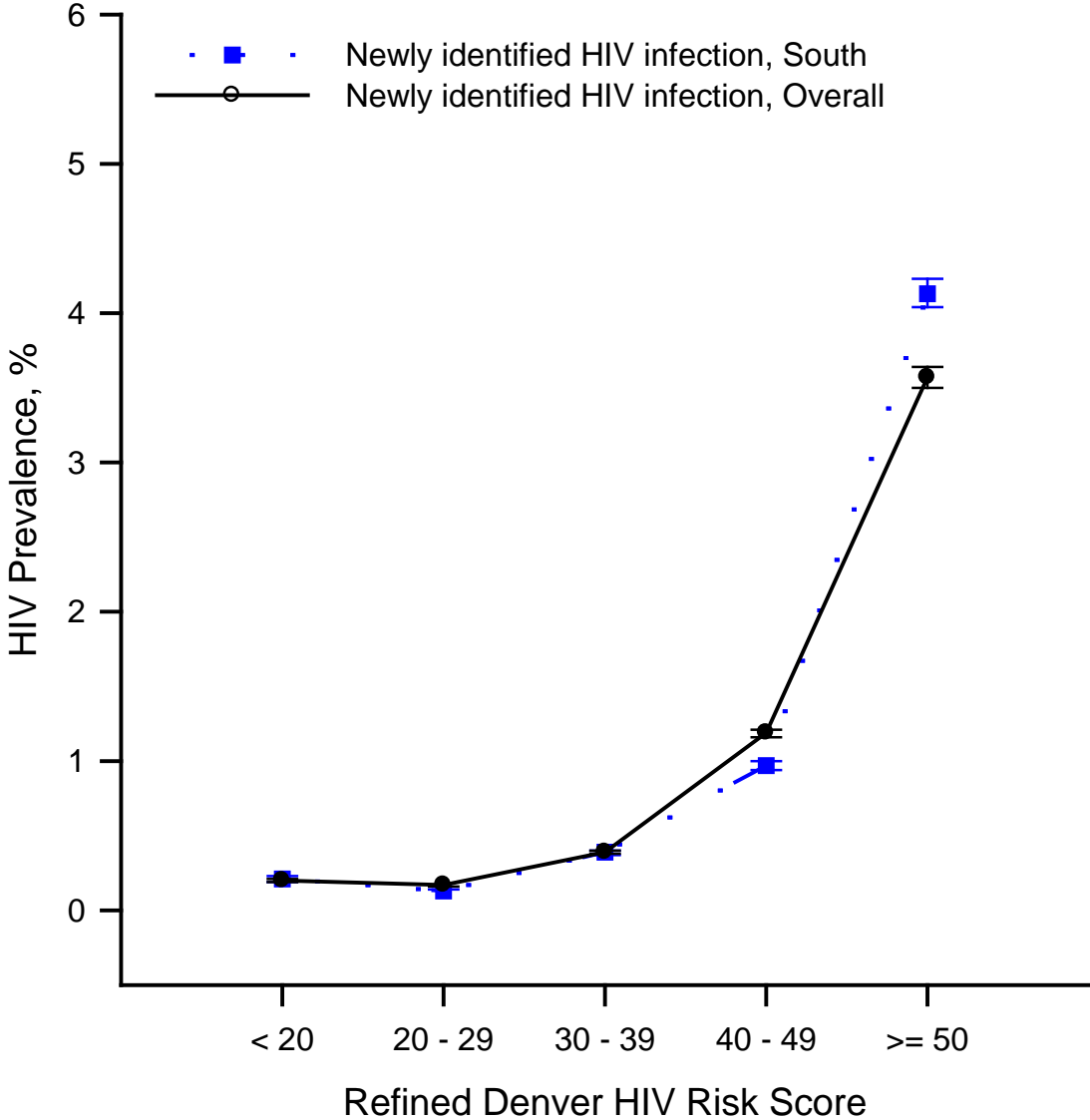|			
≥ 50	2.45	0.90 – 5.25	8.57	2.38 – 12.80

Abbreviations: CDC = Centers for Disease Control and Prevention; HIV = human immunodeficiency virus; PEMS = Program Evaluation and Monitoring System; USVI = United States Virgin Islands.
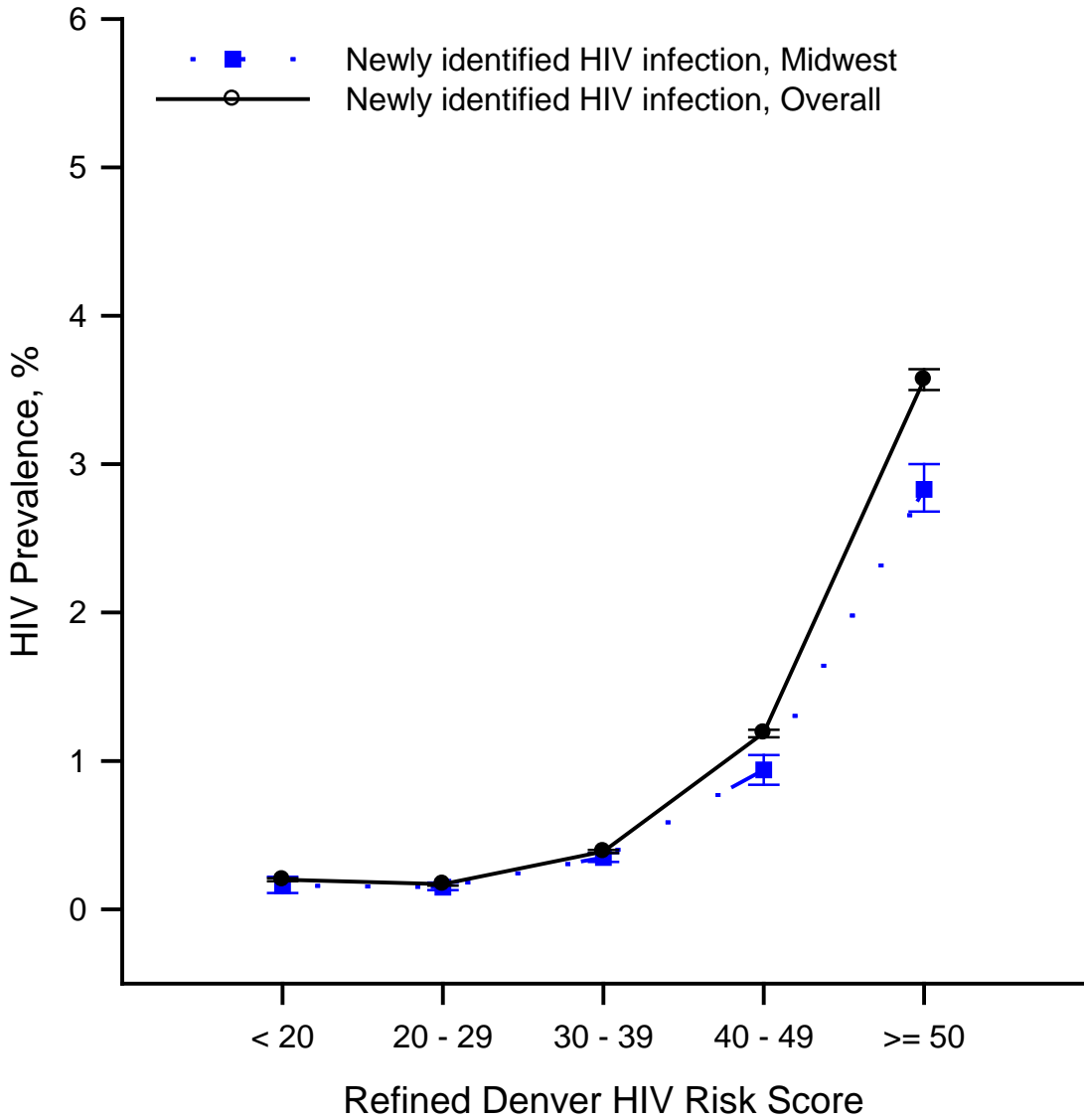
SUPPLEMENTAL Figure 2. Prevalence of newly-identified human immundeficiency virus (HIV) infection within each risk score category in the Northeast region. The regionally refined Denver HIV Risk Score ranges from -4 to +71. Bars, 95% confidence interval.
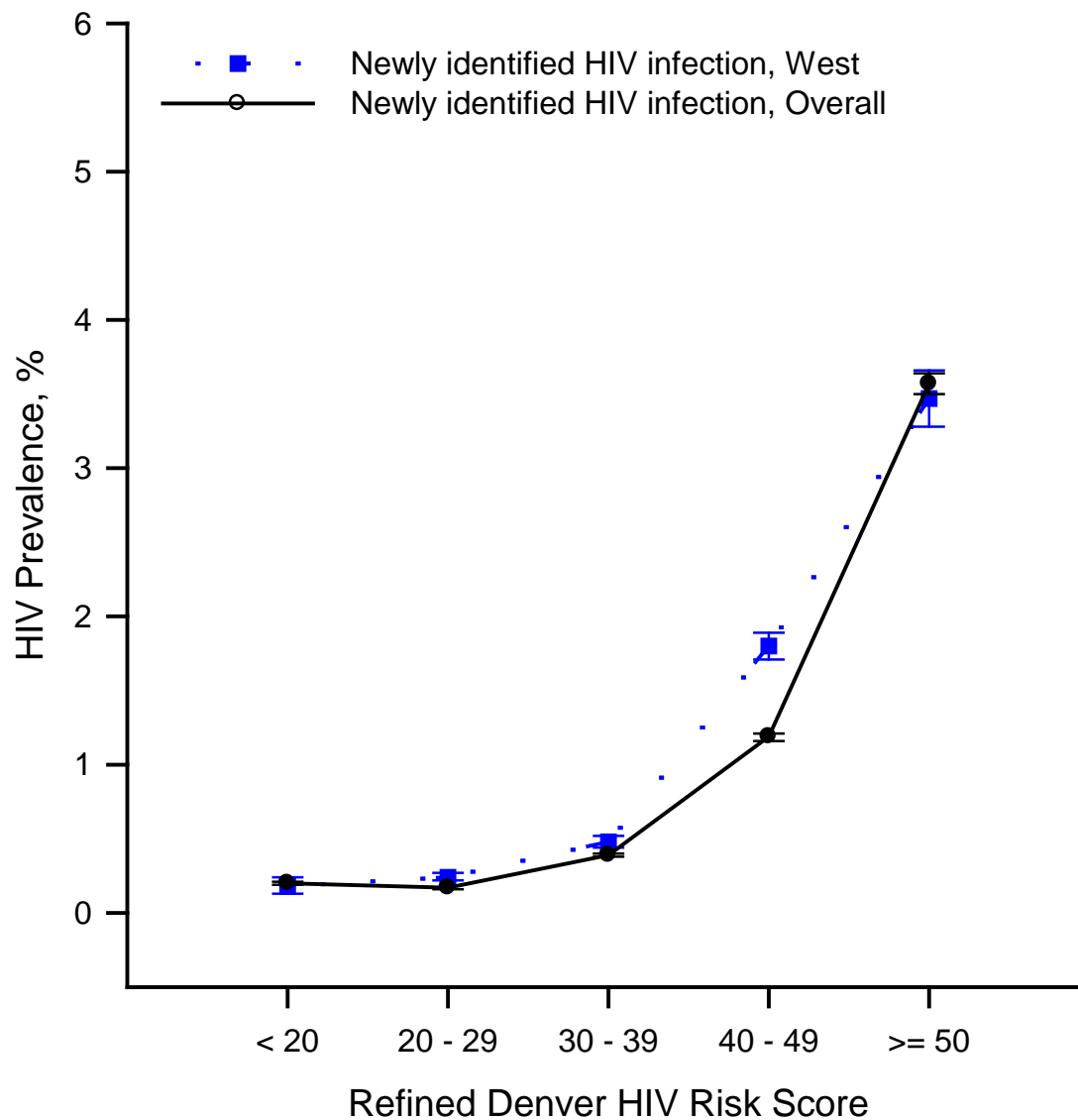
SUPPLEMENTAL Figure 3. Prevalence of newly-identified human immundeficiency virus (HIV) infection within each risk score category in the South region. The regionally refined Denver HIV Risk Score ranges from -3 to +71. Bars, 95% confidence interval.
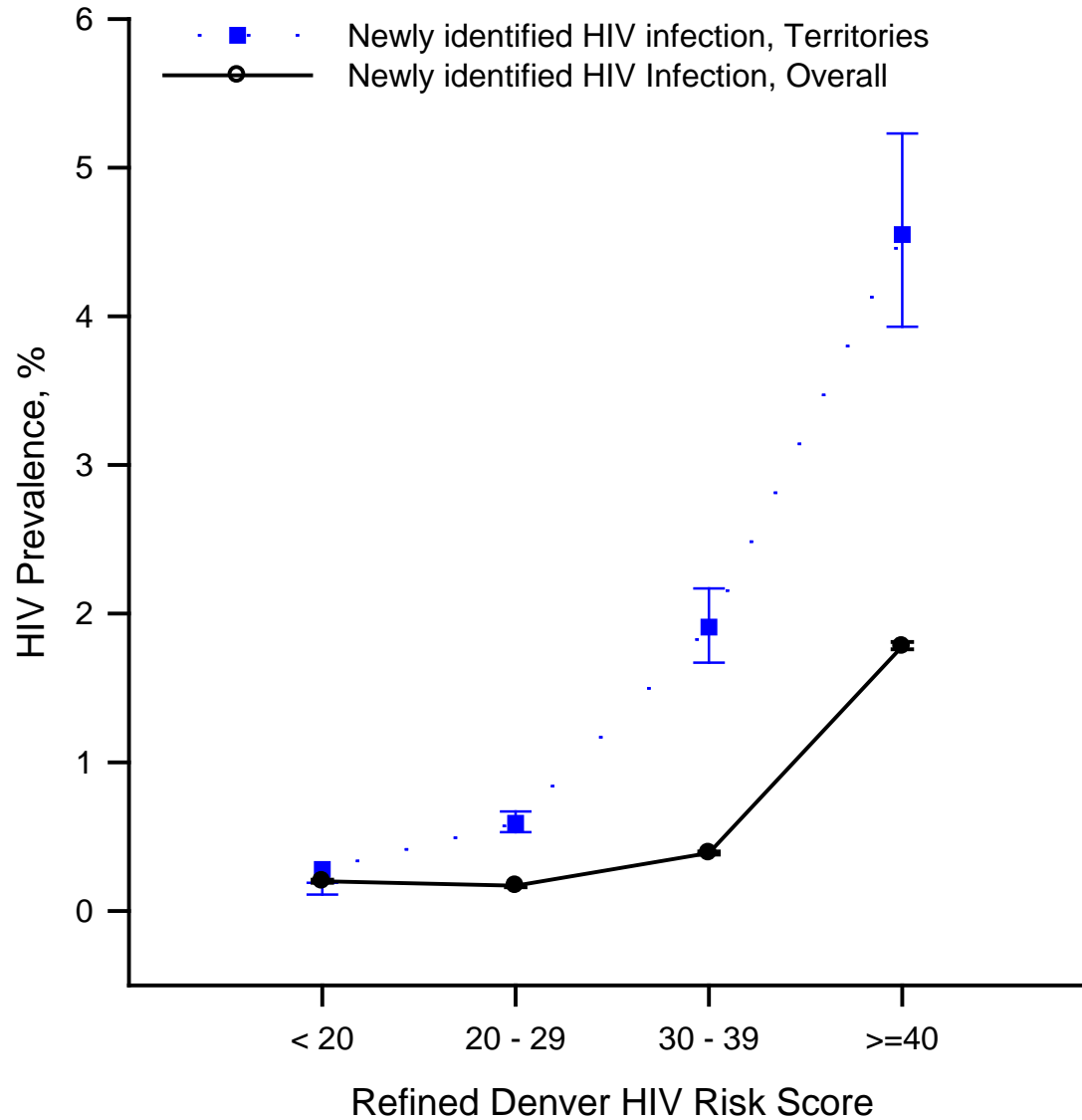
SUPPLEMENTAL Figure 4. Prevalence of newly-identified human immundeficiency virus (HIV) infection within each risk score category in the Midwest region. The regionally refined Denver HIV Risk Score ranges from 0 to +63. Bars, 95% confidence interval.

SUPPLEMENTAL Figure 5. Prevalence of newly-identified human immundeficiency virus (HIV) infection within each risk score category in the West region. The regionally refined Denver HIV Risk Score ranges from -2 to +61. Bars, 95% confidence interval.

SUPPLMENTAL Figure 6. Prevalence of newly-identified human immundeficiency virus (HIV) infection within each risk score category in the Puerto Rico and the United States Virgin Islands. The regionally refined Denver HIV Risk Score ranges from -3 to +55. Bars, 95% confidence interval.

SUPPLEMENTAL Table 7. Risk score variables for the entire cohort (N=6,768,987), stratified by HIV diagnosis, CDC PEMS data, 2008-2010.

Variable	Newly-Diagnosed HIV Infection		Confirmed HIV Test Results	
	Yes **n=43,234**	**No** **n=6,725,753**	**Positive** **n=60,230**	**Negative** **n=6,708,757**
Age				
<22 or >60 years	5,423 (12.5)	1,588,076 (23.6)	6,922 (11.5)	1,586,577 (23.7)
22-25 or 55-60 years	8,174 (18.9)	1,413,068 (21.0)	10,676 (17.7)	1,410,566 (21.0)
26-32 or 47-54 years	14,962 (34.6)	2,078,364 (30.9)	20,923 (34.7)	2,072,403 (30.9)
33-46 years	14,498 (33.5)	1,615,518 (24.0)	21,479 (35.7)	1,608,537 (24.0)
Missing	*177 (0.4)*	*30,727 (0.5)*	*230 (0.4)*	*30,674 (0.5)*
Gender				
Female	10,105 (23.4)	3,326,988 (49.5)	15,192 (25.2)	3,321,901 (49.5)
Male	32,672 (75.6)	3,352,048 (49.8)	44,415 (73.7)	3,340,305 (49.8)
Missing	*497 (1.1)*	*46,717 (0.7)*	*623 (1.0)*	*46,551 (0.7)*
Race/Ethnicity				
Black	23,936 (55.4)	3,002,307 (44.6)	32,939 (54.7)	2,993,304 (44.6)
Hispanic	7,920 (18.3)	1,338,159 (19.9)	11,494 (19.1)	1,334,585 (19.9)
Other* or White	10,003 (23.4)	2,124,985 (31.6)	13,980 (23.2)	2,121,008 (31.6)
Missing	*1,375 (3.2)*	*26,032 (3.9)*	*1,817 (3.0)*	*259,860 (3.9)*
Sexual Practices				
Sex with a male	24,227 (56.0)	2,983,515 (44.4)	33,046 (54.9)	2,974,696 (44.3)
Missing	*4,675 (10.8)*	*833,117 (12.4)*	*5,700 (9.5)*	*832,092 (12.4)*
Other Risks				
Injection drug use	20,660 (4.8)	208,259 (3.1)	3,107 (5.2)	207,218 (3.1)
Missing	*6,480 (15.0)*	*1,029,798 (15.3)*	*8,117 (13.5)*	*1,028,161 (15.3)*
Past HIV test	24,760 (57.3)	3,776,298 (56.2)	41,071 (68.2)	3,759,987 (56.1)
Missing	*9,043 (20.1)*	*1,146,392 (17.0)*	*9,331 (15.5)*	*1,146,104 (17.1)*

Percentages do not always add to 100% due to rounding error.

Abbreviations: CDC = Centers for Disease Control and Prevention; HIV = human immunodeficiency virus; PEMS = Program Evaluation and Monitoring System.

*Represents American or Alaskan Native, Native Hawaiian, or non-Hawaiian Pacific Islander.

SUPPLEMENTAL Table 8. Complete case (N=4,830,941) results and best-case/worst-case sensitivity analyses (N=6,768,987), CDC PEMS data, 2008-2010.
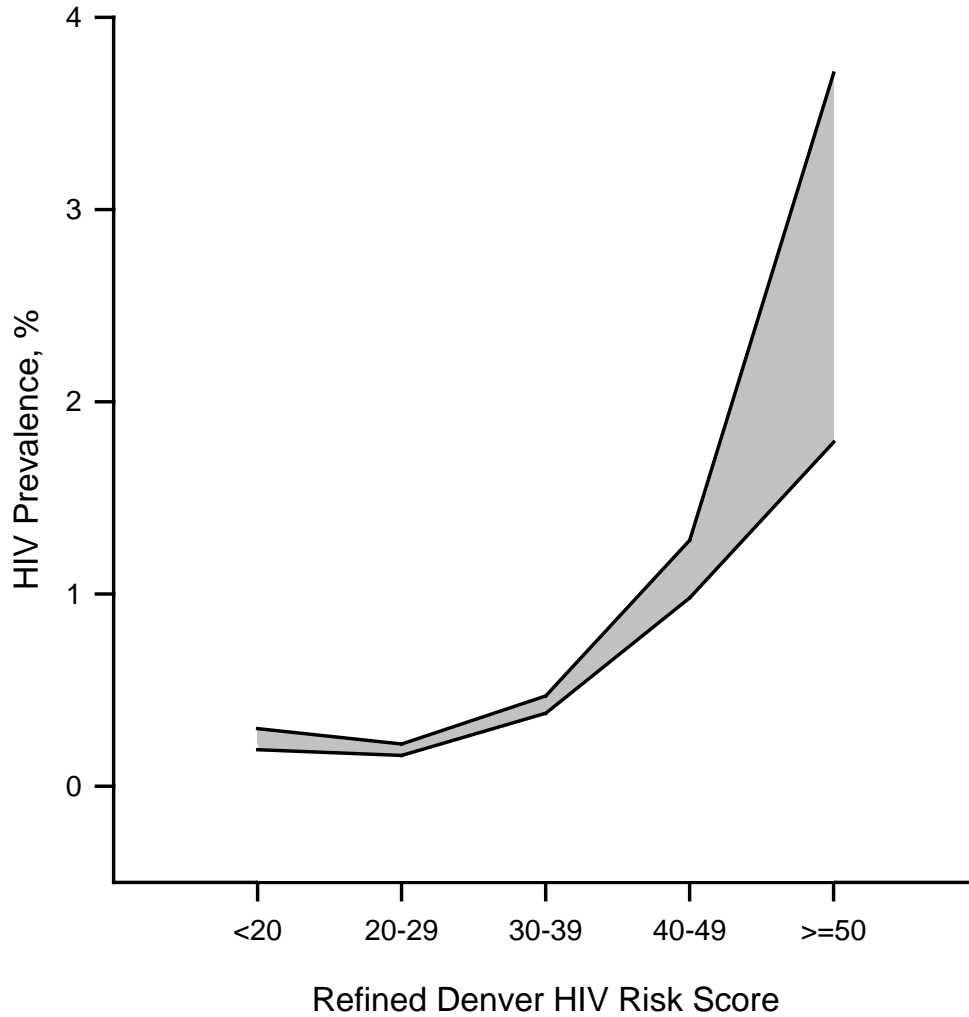
DHRS	Newly-Diagnosed HIV Infection		Confirmed HIV Test Results	
	Prevalence, %	(95 % CI)	Prevalence, %	(95 % CI)
Complete Case				
< 20	0.20	(0.19 – 0.21)	0.50	(0.48 – 0.52)
20 – 29	0.17	(0.16 – 0.17)	0.28	(0.27 – 0.29)
30 – 39	0.39	(0.38 – 0.40)	0.62	(0.61 – 0.63)
40 – 49	1.19	(1.16 – 1.21)	1.64	(1.61 – 1.66)
≥ 50	3.57	(3.50 – 3.64)	4.79	(4.71 – 4.87)
Sensitivity Analysis I*				
< 20	0.28	(0.27 – 0.30)	0.49	(0.47 – 0.50)
20 – 29	0.19	(0.18 – 0.20)	0.29	(0.29 – 0.30)
30 – 39	0.42	(0.41 – 0.42)	0.62	(0.61 – 0.63)
40 – 49	1.00	(0.98 – 1.01)	1.34	(1.32 – 1.37)
≥ 50	1.82	(1.79 – 1.85)	2.37	(2.33 – 2.40)
Sensitivity Analysis II†				
< 20	0.25	(0.24 – 0.26)	0.41	(0.40 – 0.42)
20 – 29	0.22	(0.21 – 0.22)	0.33	(0.32 – 0.34)
30 – 39	0.46	(0.45 – 0.47)	0.67	(0.66 – 0.68)
40 – 49	1.26	(1.23 – 1.28)	1.64	(1.62 – 1.67)
≥ 50	3.64	(3.58 – 3.71)	4.82	(4.74 – 4.89)

Abbreviations: DHRS = Denver HIV Risk Score; HIV = human immunodeficiency virus; PEMS = Program Evaluation and Monitoring System.
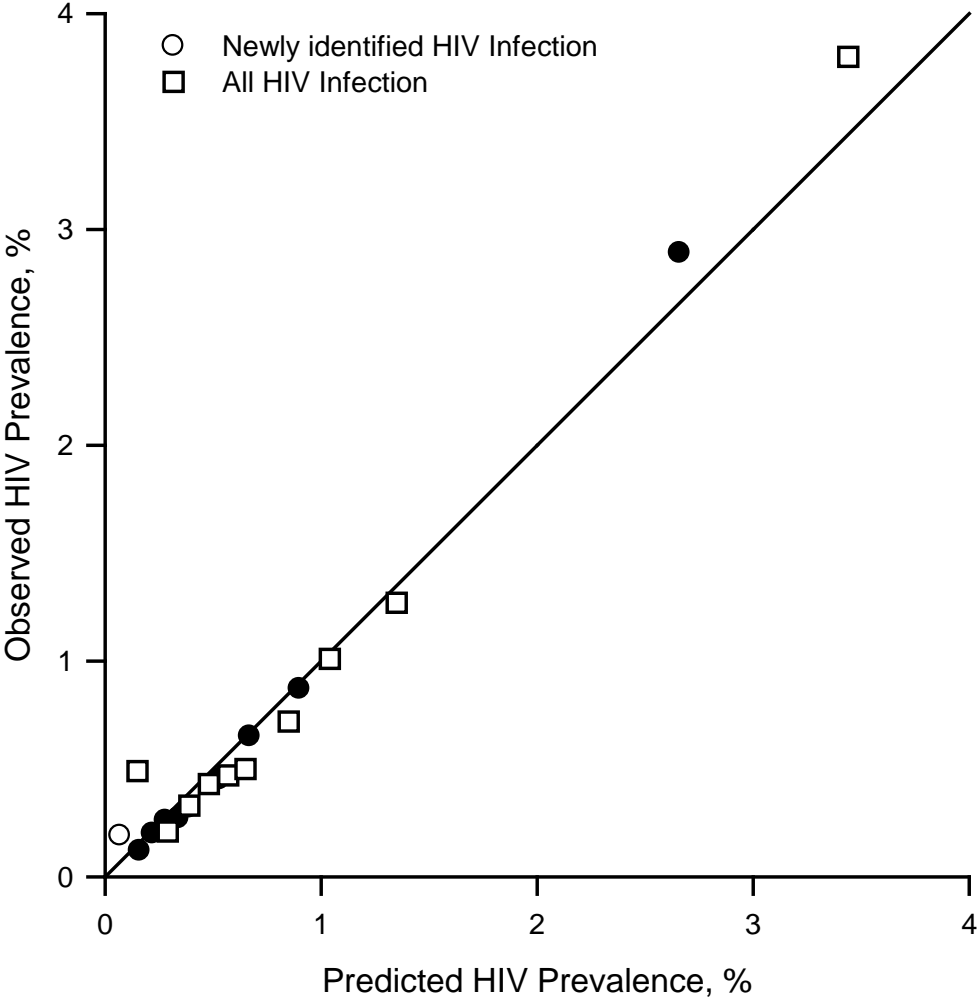
* All missing values were replaced with values to generate the highest possible DHRS (e.g., sex with a male was coded as +22; past HIV test was coded as -4).

† All missing values were replaced with values to generate the lowest possible DHRS (i.e., all missing values were coded as zero).

SUPPLEMENTAL Figure 7. Results of the sensitivity analyses showing the prevalence of newly-identified human immunodeficiency virus (HIV) infection for each risk score category of the refined DHRS when replacing all missing values with values that represented highest and lowest possible scores. The top and bottom lines represent the most extreme 95% confidence limits of the prevalence estimates, while the shaded region represents the possible range of the performance of the DHRS.

SUPPLMENTAL Figure 8. Calibration of the refined Denver HIV Risk Score to identify patients at risk of undiagnosed and all human immunodeficiency virus (HIV) infection, CDC PEMS Data, 2008 - 2010.

SUPPLEMENTAL Figure 9. Discrimination of the refined Denver HIV Risk Score to identify patients at risk for undiagnosed human immunodeficiency virus (HIV) and all HIV infection, CDC PEMS data, 2008 - 2010. The area under the receiver operating characteristic curve was 0.77 (95% confidence interval: 0.77, 0.77) and 0.74 (95% confidence interval: 0.73, 0.74) for newly identified HIV infections and all HIV infections, respectively.