

Proposed Supplement to leDEA Region 2: Caribbean, Central and South America network for HIV research (CCASAnet) for cancer-related studies

Research Plan

A. Specific Aims

1. To establish a region-wide registry of cancer-related information on approximately 12,000 HIV positive individuals in six countries in Latin America
2. Use the registry to analyze and describe the natural history of treated and untreated cancer in the setting of HIV infection in the region
3. Undertake pilot studies to characterize observed differences in patterns of cancer diagnosis, treatment and outcomes observed in the region
4. Develop data analysis tools to facilitate creation and analysis of the registry

B. Background and Significance

Effective antiretroviral therapy is changing the face of the global HIV epidemic in ways that have importance for cancer incidence and outcomes. HIV-1 infection predisposes to the development of specific types of cancer. Most cancers seen in the AIDS setting are related to oncogenic virus infections, such as Epstein-Barr virus (EBV), Kaposi's sarcoma (KS)-associated herpesvirus (KSHV) and human papillomavirus (HPV). It is generally assumed that HIV-1 infection plays a passive role in cancer development by impairing the host immune surveillance and increasing the risk of oncogenic virus infection. Recent insights, however, indicate that HIV-1 infection more actively promotes cancer growth. [Aoki 2004, Engels 2007]. Consistent with this, in a recent report from the Swiss HIV cohort, statistically significant elevated incidence ratios were observed for anal cancer, Hodgkin lymphomas, cancers of the cervix, liver, lip, mouth, pharynx, trachea, lung, bronchus and skin (nonmelanoma). In HAART users, standardized incidence ratios for KS and non-Hodgkin lymphoma were lower than those for not on HAART. [Clifford 2005] Comparable data is not available for the Latin American region, which has a wide range of socioeconomic conditions, environmental conditions and ethnic/genetic influences that differ from the reported European cohort.

The Caribbean, Central and South America network for HIV epidemiology (CCASAnet) has created a resource of clinical data on more than 12000 HIV-infected individuals in Argentina, Brazil, Chile, Haiti, Honduras and Peru. Characteristics of the patient populations and the data resource is shown in Table 1

Table 1. Program characteristics of sites participating in CCASAnet [adapted from McGowan 2007].

Site	Argentina	Brazil	Chile	Haiti	Honduras	Peru
	Fundación Huésped Buenos Aires	Projeto Praça Onze Rio de Janeiro	Fundación Arriarán Santiago	Les Centres GHESKIO Port-au-Prince	Instituto Hondureño de Seguridad Social and Hospital Escuela Tegucigalpa	Universidad Peruana Cayetano Heredia Lima
Total no. of adult patients	2000	2800	1450	4000	1000	710
No. of adult patients on ART	1600	2500	1150	4000	750	500
Measure of ART adherence	Patient self-report	Electronic registry of ART pick-up (monthly)	Electronic registry of ART pick-up (monthly)	Electronic registry of ART pick-up (monthly), pill counts, patient self-report	Registry of ART pick-up (twice monthly), patient self-report	Registry of ART pick-up (twice monthly)
Use of virologic monitoring	Yes	Yes	Yes	No	Yes (IHSS) Few (HE)	Yes
Viral genotyping/ subtyping	Yes (few)/Yes	Yes/Yes	Yes/No	No/No	Yes (few)/No	No/No
Biological specimen archive	Yes	Yes	Yes (100)	No	No	Yes (300)
Active tracking of loss to follow up/percent loss to follow up (%)	No/20	Yes/6	No/8	Yes/8	Yes/10 (IHSS) No/20 (HE)	Yes/25

Institutional Review Board	Local	Local and federal	Local	Local, national, Cornell, and Vanderbilt	Local	Local
----------------------------	-------	-------------------	-------	--	-------	-------

ART, antiretroviral therapy; HAART, highly active ART; MOH, Ministry of Health; GFATM, The Global Fund to Fight AIDS, Tuberculosis and Malaria; PHS, Public Health Service; PEPFAR, U.S. President's Emergency Plan for AIDS Relief; NIH, U.S. National Institutes of Health; NRTI, nucleoside reverse transcriptase inhibitor; NVP, nevirapine; SQV, saquinavir; RTV, ritonavir; WHO, World Health Organization; IHSS, Instituto Hondureño de Seguridad Social; HE, Hospital Escuela. Costs reported in U.S. dollars. Approximate number of samples in given category is included in parentheses unless otherwise indicated.

^awww.sadi.org.ar/images/RecomendacionesTAARV_SADI_2006.pdf

^b<http://www.aids.gov.br>

^chttp://www.minsal.cl/ici/guiasclinicas/vihsidaR_Mayo10.pdf

^d as recommended by the Haitian government

^ewww.minsa.gob.pe

Core aims of CCASAnet are (i) to create and support a network of participating sites in the Caribbean and Central and South America for sharing of existing research and clinical data related to the epidemiology of HIV and related disorders; (ii) to create a shared data repository and associated technologies for data merging that forms the union of data sets submitted by sites; (iii) to conduct and facilitate research using the shared data repository that enables answers to questions that cannot be answered by any single source; (iv) to develop and evaluate new biostatistical methods relevant to HIV epidemiology; (v) to develop a program of education and training that will assist sites to improve the quality and consistency of their clinical research activities; and (vi) to participate with other regional IEDEA networks in the development of international standards for sharing and meta-analysis of HIV-related data.

C. Preliminary Studies

Characterization of cancer in the region

Members of the CCASAnet network have participated in the characterization of cancer incidence, clinical presentation, and treatment of cancers in the region. Verdonck, Gotuzzo and colleagues from CCASAnet's site in Peru recently published a review of HTLV1-associated neoplastic diseases [Verdonck 2007]. They have also characterized the Peruvian form of KS [Mohanna 2005a, 2005b] and determined that Peru has a unique strain of HHV8 (E subtype) capable of producing KS in indigenous Peruvian persons, differing from Brasil and Ecuador where the same strain does not produce symptoms. [Mohanna 2007].

In Brazil, Mauro Schechter and colleagues have found that the seroprevalence of antibodies to HHV8 in Kaposi Sarcoma patients in Brazil was similar to that found in KS patients in the US [Zhang 1998], and determined the prevalence of HPV in women with HIV [Grinsztejn 2006]. They have also characterized the clinical manifestations of HTLV1 in Brazil, including an unusual form of myelopathy and other types of neurologic complications [Cavalcanti 1993, Harrison 1997a], the incidence of HTLV1 and HIV seropositivity among healthy blood donors [Nogueira 1996], and the clinical outcomes of HTLV1 and HIV co-infected individuals [Schechter 1994, Schechter 1997, Harrison 1997b, Harrison 1998].

Cancer-related data

The data currently acquired by CCASAnet sites is listed in Table 2. AIDS-related cancers such as KS and NHL are currently recorded within the cohort, and other forms of cancer diagnosis are recorded in the clinical databases at some but not all sites.

Table 2. Available data elements for patients included in CCASAnet consortium [McGowan 2007]

	Argentina	Brazil	Chile	Haiti	Honduras	Peru
Demographics (birthdate, gender)	Yes	Yes	Yes	Yes	Yes	Yes
Education level	Yes	Yes	Yes	Yes	No	Yes
Household income	No	No	No	Yes	Yes	No
Risk behaviors (condom use, IDU, MTM, CSW)	Yes	Yes	Yes	No	Yes	Yes

Weight at diagnosis	Yes	Yes	Yes	Yes	Yes	Yes
Weight follow-up	Yes	Yes	Yes	Yes	Yes	Yes
Height	Yes	Yes	No, but can be captured prospectively	No		Yes
Specific OI diagnoses	Yes	Yes	Yes	Yes	Yes	Yes
Specific STD diagnoses	Yes	Yes	Yes	No	Yes	Yes
Specific non-AIDS related diagnoses	Yes	Yes (after 2005)	Yes	Yes		No
Cause of death	Yes	Yes	Yes	Yes	Yes	Yes
Dates of ART	Yes	Yes	Yes	Yes	Yes	Yes
Reasons for ART discontinuation	Yes	Yes	Yes	Yes	Yes	Yes
Medications other than ART	Yes	Yes	Yes	Yes	Yes	No
CD4 ⁺ lymphocyte count (baseline/follow-up)	Yes/Yes	Yes/Yes	Yes/Yes	Yes/ Yes	Yes/ Yes	Yes/Yes
Plasma HIV-1 RNA level (baseline/follow-up)	Yes/Yes	Yes/Yes	Yes/Yes	No/No	Yes/ Yes	Yes/Yes
Hemogram	Yes	Yes	Yes	Yes	Yes	Yes
HIV genotype	Yes (few)	Yes	Yes (after 1 st ARV failure)	No	Yes	No
HIV subtype	Yes	Yes	No	No	No	No
Liver function tests	Yes	Yes	Yes	Yes	Yes	Yes
Serum creatinine	Yes	Yes	Yes	Yes	Yes	Yes
Lipids	Yes	Yes	Yes	No	Yes	Yes
Toxoplasma serology	Yes	Yes	Yes	No	Yes	Yes
Hepatitis B serology	Yes	Yes	Yes	No	Yes	Yes
Hepatitis C serology	Yes	Yes	Yes (few)	No	No	Yes
Syphilis serology	Yes	Yes	Yes	Yes	Yes	Yes
Pap smear	Yes (few)	No	Yes (few)	No	No	No

IDU, injection drug use; MTM, male to male transmission; CSW, commercial sex worker; ART, antiretroviral therapy; OI, opportunistic infection; STD, sexually transmitted disease.

The methods used to acquire these data for related to HIV care can be applied also to acquiring cancer-related clinical information from participating sites in the region.

Database linkage

A key requirement for linking HIV, cancer and mortality is the ability to link records from participants in the current cohorts to national cancer registries and death index databases. A national person unique identifier exists in all countries represented in the consortium with the important exception of Brazil (population 180 million). In Brazil, key health-related registry data is available for research, but is recorded only by person name, mother's name and date of birth. As a pilot test a computerized algorithm was developed by Antonio Guilherme of Projeto Praça Onze in Rio de Janeiro to perform linkage between electronic databases. The variables that were used for the linkage include name, mother's name and date of birth. A phonetic algorithm (PA) adapted from the original Soundex algorithm to account for Portuguese/Brazilian names was developed to yield appropriate phonetic codes and used to perform a blocking of the first and last names (both patient's and mother's) codes; a combination of these codes was used as minimum criterion for two records to be considered similar.

Preliminary analysis of sensitivity and specificity of the algorithm using data from two large HIV-infected cohorts in Rio de Janeiro and the municipality all-cause mortality database showed that it reaches high sensitivity and specificity for records with the three variables available and constitutes a valuable tool to recover patients lost to follow up in cohorts, as shown in Table. The lack of mother's name did not affect its accuracy, but records with patient's name only have much lower sensitivity and specificity, probably due to the high prevalence of patients with identical names in Brazil. Thus the algorithm is expected to work well to link cancer registry and death index databases that have at least patient's names and dates of birth.

Table 3: Matching algorithm performance for national cause of death registry matched to UFRJ research data records

Data Elements Available	Match Sensitivity [95% C.I.]	Match Specificity [95% C.I.]
Patient & Mothers Name, Date of Birth	93.9 [88.1, 99.7]	99.5% [98.5, 100]
Patient Name, Date of Birth	92.1 [86.0, 98.1]	99.5% [98.5, 100]
Patient Name Only	58.5 [43.4, 73.6]	88.5% [84.1, 92.9]

D. Research Design and Methods

The work proposed here is an enhancement to activities already underway in leDEA Region 2 under NIH Cooperative agreement 1 U01 AI069923. The overall effort will be directed by the CCASAnet principal investigator, Daniel Masys MD, who is a board-certified medical oncologist and former chief of the NCI's International Cancer Research Data Bank branch [Hubbard 1986, Masys 1987].

We propose four specific aims:

Aim 1: To establish a region-wide registry of cancer-related information on approximately 12,000 HIV positive individuals in six countries in Latin America

Using the patient records already available in the CCASAnet data sources, each site will develop a strategy to acquire information related to cancer diagnosis, treatment and outcome. This information will come from a variety of sources, including review and coding of paper charts using ICD9/10 oncology diagnoses, and linkage of existing databases to regional and national tumor registries. All cancers, not just cancers normally associated with HIV, will be recorded and coded, along with clinical information such as pathology reports and records of treatment. For deceased patients, attempts will be made to find corresponding records in death registries (see Aim 4) that contain diagnosis information relative to cancer and other co-morbidities beyond HIV. Because the majority of patients who are in the CCASAnet 'cohort' are under active care by the organizations listed in this proposal, clinical care records and registry records are available to CCASAnet investigators for both clinical care and research purposes. Database structures of the participating sites will be modified as needed to add variables for diagnosis codes related to cancer, cancer-related treatments, and other cancer-related events (e.g., relapse, second malignancies).

For this aim, we reasonably expect that sites will use the majority of the resources provided by the supplement to hire and train workers to perform chart reviews.

Aim 2: Use the registry to analyze and describe the natural history of treated and untreated cancer in the setting of HIV infection in the region

Once constructed and combined with the existing clinical variables available in the data collections contributing to CCASAnet, data will be submitted to the data coordination center at Vanderbilt for combined analyses. A common data dictionary will be created using methods developed by the CCASAnet data center for the merging of heterogeneous data types [Duda 2007]. These will include descriptive statistics of the numbers and types of cancers observed in the cohort, Kaplan-Meier analyses of time-to-event for key measures such as interval between HIV diagnosis and cancer diagnosis, the incidence of specific types of cancer before and after HAART initiation, relapse-free survival and overall survival. Trends in cancer diagnosis over time will be determined, and correlated with the impact of HAART therapies where feasible. Because this project necessarily makes use of existing data rather than prospectively capturing a uniform set of variables, we will carefully evaluate the effects of reporting biases that may be caused by fragmentation of health care delivery at sites and resulting inaccessibility of data related to care provided by other organizations. We will also perform analyses to examine risk factors for specific cancer types wherever feasible.

The result will be one or more publications summarizing the natural history of cancer occurring in the setting of HIV in Latin America, and trends in observed cancer incidence over time.

Aim 3: Undertake pilot studies to characterize observed differences in patterns of cancer diagnosis, treatment and outcomes observed in the region

As noted in Preliminary Studies above, there are genetic, environmental and cultural differences among people in the region. Peruvian Indians, African genetic heritage in Brazil, and European genetic heritage in Argentina may play a role in observed patterns of cancer occurring in the setting of HIV. To the extent that unusual patterns are found, more detailed studies to characterize differences in incidence, prevalence and/or cancer-related outcomes will be designed, though we appreciate that the limited funds available will not support in depth characterization of such differences using support from the proposed supplement.

Aim 4: Develop data analysis tools to facilitate creation and analysis of the registry

As an extension of the validity of the matching algorithm described in the preliminary studies section above, we propose using cancer-related registry information to further generalize its ability to perform linkage between other data sources and also to help cohorts in other settings to take advantage of the algorithm for their own needs. Minimal information required for this task would be the presence of electronic databases (e.g. a National Cancer Database and a National Mortality Database) that contains patient's name and dates of birth.

In summary, we believe that the existing Latin America cohort that is 12000 patients and increasing in size will be a powerful tool to track and analyze all types of cancer occurring in the setting of HIV infection, and will provide important insights into the evolving story of long term health risks of neoplasia that accompany antiretroviral therapy in immunocompromised individuals in widely varying socioeconomic, cultural and ethnic settings.

E. Human Subjects Research

All CCASAnet sites have IRB approval for their use and analysis of existing clinical data. Amendments to the existing IRB approvals to add cancer-related data sources will be obtained where necessary. Data is de-identified prior to submission to the Vanderbilt data coordination center for combined analysis.

G. Literature Cited

- Aoki 2004 Aoki Y, Tosato G. Neoplastic conditions in the context of HIV-1 infection. *Curr HIV Res.* 2004 Oct;2(4):343-9.
- Cavalcanti 1993 Cavalcanti, M; Ferreira-Jr, O; Puccioni, M; Novis, SAP; Schechter, M. HTLV I associated neurological manifestations in 4 generations of a Brazilian family. *J Acq Immun Def Synd* 1993; 6 (2): 213-217.
- Clifford 2005 Clifford GM, Polesel J, Rickenbach M, Dal Maso L, Keiser O, Kofler A, Rapiti E, Levi F, Jundt G, Fisch T, Bordoni A, De Weck D, Franceschi S; Swiss HIV Cohort. Cancer risk in the Swiss HIV Cohort Study: associations with immunodeficiency, smoking, and highly active antiretroviral therapy. *J Natl Cancer Inst.* 2005 Mar 16;97(6):425-32.
- Duda 2007 Duda SN, Cushman C, Masys DR. An XML Model of an Enhanced Data Dictionary to Facilitate the Exchange of Clinical Research Data in International Studies. *MEDINFO 2007 Brisbane Proceedings. Int J. Medical Informatics. Suppl.* In press
- Engels 2007 Engels EA. Infectious Agents as Causes of Non-Hodgkin Lymphoma. *Cancer Epidemiolo Biomarkers Prev* 2007;16(3), March 2007.
- Grinsztejn 2006 Grinsztejn B; Bastos FI; Veloso VG; Friedman RK; Pilotto JH; Schechter M; Derrico M; Andrade A; Lourenço MC; Moreira RI; Russomano F; Morgado M; Currier JS. Assessing sexually transmitted infections in a cohort of women living with HIV/AIDS, in Rio de Janeiro,

- Brazil. *Int J STD AIDS* 2006; 17: 473-478.
- Harrison 1997a Harrison, LH; Vaz, B; Taveira, DM; Quinn, TC; Gibbs, CJ; Souza, SH; McArthur, JC; Schechter, M. Myelopathy Among Brazilians Coinfected with Human T-Cell Lymphotropic Virus Type I and HIV. *Neurology* 1997; 48:13-17.
- Harrison 1997b Harrison, LH; Quinn, TC; Schechter, M. Human T-Cell Lymphotropic Virus Type I Does Not Increase HIV Viral Load *in Vivo*. *J Infect Dis* 1997; 175:438-440.
- Harrison 1998 Harrison, LH; Schechter, M. Coinfection with HTLV-I and HIV: Increase in HTLV-I-related outcomes but not accelerated HIV disease progression? *AIDS Patient Care and STDs* 1998; 12:619-623.
- Hubbard 1986 Hubbard, SM, Martin, NB, Blankenbaker, LW, Esterhay, RJ, Masys, DR, Tingley, DE, Stram, MC, DeVita, VT. The Physician Data Query (PDQ) cancer information system. *J Cancer Educ*, 1(2), 79-87, 1986.
- Masys 1987 Masys, DR, Hubbard, SM. Technical information programs of the National Cancer Institute. *J Am Soc Inf Sci*, 38(1), 60-4, 1987.
- McGowan 2007 McGowan C, Cahn P, Gotuzzo E, Padgett D, Pape JW, Wolff M, Schechter M, Masys D. Cohort Profile: Caribbean, Central and South America Network for HIV Research (CCASAnet) Collaboration within the International Epidemiologic Databases to Evaluate AIDS (IeDEA) Program. *Intl J Epid*, in press.
- Mohanna 2005a Mohanna S, Maco V, Bravo F, Gotuzzo E. Epidemiology and Clinical Characteristics of Classic Kaposi's Sarcoma, Seroprevalence, and Variants of Human Herpes Virus 8 in South America: A Critical Review of an Old Disease. *Int J Infect Dis* 2005; 9(5):239-250.
- Mohanna 2005b Mohanna S, Ferrufino JC, Sánchez J, Bravo F, Gotuzzo E. Epidemiological and clinical characteristics of Classic Kaposi's Sarcoma in Peru. *J Am Acad Dermatol* 2005; 53(3):435-441.
- Mohanna 2007 Mohanna S, Portillo JA, Carriquiry G, Vidal J, Ferrufino JC, Sanchez J, Bravo F, Gotuzzo E. Human herpesvirus-8 in Peruvian blood donors: a population with hyperendemic disease? *Clin Infect Dis*. 2007 Feb 15;44(4):558-61.
- Nogueira 1996 Nogueira, CM; Cavalcanti, M; Schechter, M; Ferreira-Jr, OC. Human T-Lymphotropic Virus Type I and II in Healthy Blood Donors in Rio de Janeiro, Brazil. *Vox Sag* 1996;70:47-48.
- Schechter 1994 Schechter, M; Harrison, LH; Halsey, NA; Trade, G; Santino, M; Moulton, LH; Quinn, TC. Co-infection with Human T-Cell Lymphotropic Virus Type I and HIV in Brazil: Impact on CD4 Lymphocyte counts and Clinical Stage. *JAMA* 1994; 271:353-357.
- Schechter 1997 Schechter, M; Moulton, LH; Harrison, LH. HIV Viral Load and CD4⁺ Lymphocyte Counts in Subjects Coinfected with HTLV-1 and HIV-1. *J AIDS* 1997; 15:308-311.
- Verdonck 2007 Verdonck K, Gonzalez E, Van Dooren S, Vandamme A, Vanham G, Gotuzzo E. Human T-lymphotropic virus 1 (HTLV-1): recent knowledge about an ancient infection. *Lancet Infect Dis* 2007; 7:266-81.
- Zhang 1998 Zhang, X; Fitzpatrick, L; Campbell, TB; Badaro, R; Schechter, M; Melo, MF; Brites, C; Sampaio, DP; Schooley, RT. Comparison of the Prevalence of Antibodies to Human herpesvirus-8 (Kaposi's Sarcoma- Associated Herpesvirus) in Brazil and in the United States. *J Infect Dis* 1998;178(5):1488-1491.