COHORT PILOT STUDY REPORT:
Evaluation of the Potential for an Epidemiologic Study of the Association between Work Practices and Exposure and Chronic Kidney Disease at the Ingenio San Antonio (Chichigalpa, Nicaragua)

January 30, 2012

Prepared by:
Ann Aschengrau
Daniel Brooks
Eloesa McSorley
Alejandro Riefkohl
Kate Applebaum
Juan Jose Amador
Oriana Ramirez Rubio

Additional Contributions (alphabetical order):
Bruce Cohen
James Kaufman, MD
Rebecca Laws, MPH
Damaris Lopez
Victor Hugo Palacios
Martha Pastora
Sara Romero Perez
Daniel Weiner, MD
Michael Winter

Translated from English to Spanish by:
Irving Gongora
Bricia Trejo

Boston University School of Public Health
715 Albany St – Boston, MA 02118

TABLE OF CONTENTS

I. Introduction

In its Scoping Study report of 2009, Boston University investigators recommended a number of studies designed to provide results that are relevant to the following questions, decided upon by Dialogue Table members:

1. What are the causes of chronic renal insufficiency (CRI) in the Western Zone (Zona del Occidente) of Nicaragua – an area that includes the Ingenio San Antonio and its sugarcane plantations?
2. Is there any relationship between the practices of the Ingenio San Antonio and the causes of CRI?

One of the key studies proposed was a retrospective cohort study (defined below) of former and current workers at the Ingenio San Antonio (ISA). Although the Industrial Hygiene (IH) Report completed by the Boston University group and released in 2010 concluded that based on current scientific information, none of the practices at ISA were known to cause chronic kidney disease (CKD, commonly also referred to as chronic renal insufficiency (CRI) in Nicaragua), it did not rule out the possibility that further study could lead to new knowledge that would connect some aspect of the work practices at ISA with the occurrence of CKD. The retrospective cohort study was proposed as a key component of providing results relevant to this question, particularly during the time period when most members of ASOCHIVIDA developed CKD, while also shedding light on possible non-occupational factors.

The Scoping Study also proposed a large-scale, detailed medical record review study to: (1) assess the clinical characteristics and course of CKD, including whether the clinical evidence is more consistent with tubulo-interstitial or glomerular disease; (2) evaluate medication prescription patterns, complaints, and test results potentially related to kidney function as possible etiologic factors; and (3) identify the prevalence of conditions known to cause CKD, such as diabetes or hypertension. For efficiency purposes, we decided to address the goals of the medical record review within the context of the retrospective cohort study, which itself requires medical record review. We will return to consideration of the medical record review study in a later section of this report.

Cohort studies: prospective vs. retrospective

Before proceeding further, we wish to explain what is meant by a retrospective cohort study. A **cohort study** is typically conducted within a group of people who have variability in the exposures of interest. People are usually categorized into two or more subgroups based on their level of exposure and the amount of disease that occurs in each group over a period of time is identified and compared. For example, a cohort study to assess whether smoking causes disease might be conducted among people who are categorized into groups based on whether they currently smoke cigarettes, used to smoke but have quit, or never regularly smoked. To study occupational exposures, people who have more exposure on the job to certain agents or conditions (examples from the IH Report include pesticides, heavy metals, infectious agents, silica, and heat stress/volume depletion) would be compared to people who have little or no exposure to those particular agents or conditions.

Cohort studies can be **prospective** (studying future events) or **retrospective** (studying past events). In the context of ISA workers, a prospective cohort study might start collecting information on exposure and disease among workers starting in 2012 and continue for 3-5 years. A retrospective cohort study would collect information on exposure and disease among people who worked before 2011. Because retrospective studies look at past history, they almost

always rely on existing records to obtain the necessary information on exposure and disease. The reliance on past records is both an advantage and a disadvantage compared to prospective cohort studies. Because all the information already exists, retrospective cohort studies take less time to complete and cost less money than prospective cohort studies. However, if the information in available records is not sufficient to do a good study, it is impossible to go back and obtain more information. Therefore, a key task before starting a retrospective study is to determine whether the records will be able to provide the information needed to answer the study question(s). The main concerns typically include:

1. Availability: Do the records about a subject's exposure (e.g. work as a cane cutter) and disease (e.g., elevated creatinine) exist and can they be accessed?
2. Accuracy: Is the information in the records of good quality?
3. Relevance: Is the information contained in the records sufficient to answer the study questions? What is missing, and what problems does missing information present?
4. Linkage: If information for each person comes from different sources and records, is it possible to link these records together?

The pilot retrospective cohort study at ISA

At the Dialogue session in January 2011, participants asked that the Boston University investigators present a plan for continuing research activities on CKD at ISA and in western Nicaragua through December 2011. One of the eight tasks proposed was to conduct a Pilot Study to assess the feasibility of conducting a complete retrospective cohort study examining the relationship between work practices at ISA and CKD. In particular, the Pilot Study was designed to determine the practicality of conducting a large retrospective cohort study that would be based mainly on existing employment and medical records. Secondary goals included: (1) preparing a retrospective cohort study design if such a study was considered feasible, (2) preparing recommendations to improve the system for tracking workers with CKD, (3) preparing a basic description of medical histories from ISA hospital records, and (4) gathering information that would strengthen a grant proposal to the U.S. National Institutes of Health (NIH).

In order to achieve these goals, the Pilot Study evaluated records available from the years 1997 through 2010 and included a detailed review of 243 randomly selected ISA workers with records available during this period. The Pilot Study included the following tasks: (1) review and describe all relevant data sources at ISA, (2) select Pilot Study subjects, (3) abstract demographic information as well as work duration and intensity information from ISA employment and payroll records, (4) link work and medical records, (5) abstract information from medical records, (6) computerize, clean, and analyze the collected data, (7) conduct in-person interviews with a subsample of 10 former and current workers, and (8) review Pilot Study results to address the goals described above.

II. Pilot Study Worker Population

Individuals who work at ISA can be divided into 3 different and distinct groups based on employment status: (1) contracted workers, (2) permanent employees, and (3) temporary employees. Decisions regarding groups to consider for inclusion in a retrospective cohort study and for the pilot study had to account for differences between these groups.

Contracted Workers

Contracted workers are employed by third party companies to work at ISA. ISA started negotiating work through third party contractors starting in 1995 for various jobs such as

seeders, weeders and cane cutters, but there are no known records that are accessible and have individual level employment information for contracted workers during the period 1995-1999. Starting in 2000, limited employment information was collected for contracted cane cutters, and mandatory pre-employment screening of cane cutters for health indicators including creatinine was implemented in 2003. This information was not implemented for other contracted workers until 2005. Because information on elevated creatinine levels and basic employment is of primary concern for a cohort study, cane cutters were the only contracted workers included in the pilot study, since they were the only workers with creatinine measurements.

Permanent and Temporary Workers
Permanent and temporary employees are direct employees of ISA. Distinctions between these two categories of workers have varied over time. One key difference is in the frequency of creatinine testing. Since a policy of mandatory pre-employment creatinine screening for ISA employees (but not contracted workers) was initiated in 2000, permanent employees have been screened only at initial hire, whereas temporary employees have been screened at the beginning of each new contract, generally on at least an annual basis. Because an elevated creatinine level would be a primary outcome of a cohort study and permanent workers would not have sufficient testing to determine whether their creatinine level was elevated, they were not considered for inclusion in a retrospective cohort study and were excluded from the pilot study.

Finally, payroll information is available for temporary employees beginning in 1997, which is also the year that ISA implemented creatinine testing on a large scale. Although screening was not mandatory until 2000, it was made widely available to ISA employees. For these reasons, only workers employed in 1997 or later were considered for the pilot study. This restriction does allow for the inclusion of persons who began working before 1997 as long as they continued to work and had records available after that time.

Based on the above considerations, the pilot study was limited to workers who at the time of selection were: (1) non-cane cutter temporary workers employed sometime between July 1, 1997 and June 30, 2010, (2) cane-cutters who worked at ISA via direct employment sometime between 1997 and 2003, or (3) cane cutters who worked at ISA via a third-party contractor sometime between 2003 and 2009.

III. Review of Occupational Data Sources

In this section, we document our review of the data available at ISA which could be used in a retrospective cohort study to track individual work histories over time. An individual's employment history at ISA is documented in numerous data sources, including employment and payroll records for individuals directly employed by ISA and other sources for contracted workers. Starting in 1997, some cane cutters were hired through third party contractors, and all cane cutters were hired in this manner by 2003. The employment history of cane cutters is documented in ISA records through 2002. The employment history of cane cutters hired by contractors can be documented using various data sources including the enganche database compiled and maintained by the Office of Responsibilidad Social Salud en el Campo, Harvest Office payroll records, and Ordenes de Pegue records. The following sections provide brief descriptions of all these data sources.

A. Records of Individuals Directly Employed by ISA

1. Paper Employment Records: ISA maintains hard copies of employment files for individuals who have been employed directly by the Ingenio. These files only contain employment history information relevant to the time of direct employment by ISA. Paper employment files for workers who were ISA employees in 1990 or later are physically located in the payroll/employment building and the records of workers whose employment was terminated before 1990 are archived elsewhere. Beginning in 2006, paper employment records are based on print outs of electronic records (see section below). Files containing paper records are grouped into three categories: current employees, retirees, and former employees who have indicated that they have CKD. Within these categories, files are organized by ficha number[1], a unique employee ID number given to all ISA employees. Documents that may be in a worker's employment file include the employment history form, work contracts, pre-employment physical and health screening forms, social security registration form, proof of beneficiaries, certificate of education, and police registration. Paper copies of the employment history records includes information such as demographic characteristics, dates of employment, planillas, job titles, pay raises, and time off for vacation and illness. The term "planilla" is best described as the payroll division used for processing workers' wages .

2. Electronic Employment Records: ISA also maintains complete electronic files of its workers from 2006 and later. These files include the worker's employment history since 2006, as well as employee characteristics found in the paper employment records (described above) plus some additional information. .

3. Paper Payroll Records: Like its employment files, ISA payroll documents contain information only for workers who were employed directly by the Ingenio. As such, there are ISA payroll records for cane cutters only through the end of the 2002-2003 zafra, after which all cane cutters were contracted. The earliest paper payroll records date from July 1997. The physical condition of these records varies; some are faded and some have torn edges or pages. The payroll office does not handle payroll information for Planillas 01, 03, 06, 07, 08 or 09, but workers in these Planillas are predominantly permanent workers (e.g., engineers and executives) and are not included in the Pilot Study. The payroll office processes hand payments for Planillas 02 – office based supervisors, 04 – machine operators, 05 – field workers, 10and 11 – cane cutters, and 12 and 13 – temporary factory workers.

The blue and white paper payroll "books" contain only partial payroll information for each 14-day period, but complete payroll information is in the electronic records (see below). Payroll records are sorted by planilla, date and "codigo proceso," a four or five digit accounting code. The planilla and codigo proceso codes are noted on the top of each page of the "book" and correspond to "proceso planilla" in the employment history record. Individual records are organized within each codigo proceso by ficha number. Information in payroll records includes the number of regular and extra hours worked, the number of days worked in each 14 day period, and details regarding vacation earned and payment. Codigos Funcionales and Codigos Procesos changed at various points over time, including a major coding change in 2001. These changes make establishing a person's work history over time complicated, and so it has been important to gain an understanding of the major code changes over time

[1] In this report, we use "ficha" to designate the ID number for ISA employees and "carnet" number for the ID number for contracted workers. In reality, they are often used interchangeably.

4. Electronic Payroll Records: Electronic ISA payroll records are available beginning in November 2000. Electronic payroll information is available on a daily basis, but the default presentation is weekly. The electronic file contains all the information in the paper payroll records and important additional information, such as the numerical "codigo funcional" (CODCAT) and "codigo proceso" (CODPRO), along with text descriptors. The former represents the job title, which corresponds to "numero puesto" in employment files before 2006, and the latter corresponds to the type or category of work. Due to inclusion of these variables, the electronic payroll records can be used to determine employment history from November 2000 forward. We have a partial listing of variables in the electronic payroll records.

B. Records of Contracted Cane Cutters

Employment histories of cane cutters can be divided into 3 periods according to the circumstances of their employment. Up until the 1997-1998 zafra, all cane cutters were directly hired by ISA and employment histories during this period are in the records described above. An increasing proportion of cane cutters were hired via contractors starting in the 1998-1999 zafra, and all cane cutters have been hired via contractors from the 2003-2004 zafra to the present time (see Table 1 below).

Table 1: Number and Percent of Cane Cutters by Zafra and Employer

	1998-99	99-00	2000-01	01-02	02-03	03-04	04-05	05-06	06-07	07-08	08-09	09-10
All Cane Cutters	1500	1520	1470	1470	1590	1350	1230	1280	1170	1270	1120	1070
Contract-ed Cane Cutters	300	420	370	670	790	1350	1230	1280	1170	1270	1120	1070
ISA Employed Cane Cutters	1200	1100	1100	800	800	0	0	0	0	0	0	0
Percent of Cane Cutters who were ISA employees	80%	72%	75%	54%	50%	0%	0%	0%	0%	0%	0%	0%

The following sections describe the various sources of employment history information that can be retrieved for contracted cane cutters during the period 2003-2010.

1. Enganche Screening Database: Beginning in 2003, mandatory screening of all potential contracted cane cutters for elevated creatinine levels was implemented by the Office of Responsibilidad Social Salud en el Campo under the direction of Dr. Felix Zelaya. Records for each enganche are maintained in separate Excel files. These files contain information on all cane cutters who were found eligible for work based on the criteria that they have a creatinine level of <1.3 mg/dL (retests were permitted for values slightly above the cut-off level) and no other specific conditions (e.g., uncontrolled hypertension or diabetes), as well as those that did not meet these criteria. However, not all eligible workers in a given year were hired to work at ISA for that zafra, and the enganche databases do not indicate which workers were eventually hired. The level of information in each enganche database varies; however, all databases

include *at least* full name, INSS and cedula numbers, if available, and the contracted worker ID (carnet).

2. Ordenes de Pegue: When workers are cutting cane, they are grouped by contractor and by the particular area of the field in which they are working. Each group of workers is known as a "pegue," and a vehicle picks up piles of cut cane which are then weighed to determine the tonnage cut per "pegue". This vehicle produces a ticket on a daily basis that contains the tonnage and the carnet number for each worker in the "pegue." Up to 16 workers are recorded electronically on the ticket; additional workers are noted in writing. Starting in 2000, the first 16 worker carnet numbers and the tonnage information have been entered into an electronic file known as the "Ordenes de Pegue". Carnet numbers from the Encanche Screening Database can be linked to the Ordenes de Pegue starting in 2003 (when the Enganche Screening Database became available) to establish employment of contracted cane cutters during this period. Duration and intensity of work as a contracted cane cutter can be assessed through the Ordenes de Pegue to the extent that there is a listing of a worker's carnet number for each day worked.

A separate database also contains information of sugar cane cut by workers from the 1999-2000 to the 2002-2003 zafras. It contains 4,356,391 entries, each with the ficha number, name and family name of the worker. It also contains the date when the cane was cut, contractor abbreviation, vehicle number, field, and the ruma (row of cut cane in the field). Finally, it contains the number of quintales cut (unit of weight equal to 45.45 kilograms or 100 pounds), the price per quintal, whether the day was a Sunday or a holiday, and the amount received. In a separate table, it is possible to find more information about each worker (a total of 27,680) by ficha number, date first ever worked for the company, date and place of birth, sex, complete name, address, social security number and cedula.

3. Contracted Cane Cutter Payroll Records: Beginning in 2007, the ISA Harvest Office began maintaining payroll records for contracted workers. These payroll records consist of PDF files organized by contractor and carnet number. Files are produced every week throughout the zafra. These records include information such as identification number, planilla, dates of work, and productivity measures. It is our understanding that these PDF files were produced from a database which contains these variables among many others.

IV. Review of Medical Information Sources

The availability of information on the medical outcomes under study is key to determining if a retrospective cohort study can be successful. It is important to note that all cohort studies begin with healthy subjects who are followed to determine who develops the disease(s) of interest. Thus, these types of studies measure the occurrence of new or incident cases of disease in a defined population over a specified time period. Incident cases of disease can be defined in a variety of ways. For example, in this setting, definitions of disease could range from elevated creatinine levels to a clinical diagnosis of CKD.

The ISA Hospital appears to be the main source of medical care for directly employed workers. However, other sources of care also need to be considered, particularly before and after employment at ISA, and so the Pilot Study included a review of records at the Centro de Salud in Chichigalpa, Hospital Escuela Oscar Danilo Rosales Arguello (HEODRA) Hospital, and a few other medical facilities and health centers in the area.

A. ISA Hospital Records: Records at the ISA Hospital are sorted by INSS number. If no INSS number is available and the worker was still employed as of October 2008, their INSS number can be searched by name and date of birth in computerized Excel files maintained by the hospital. Starting in October 2008, there is one electronic file per month that contains a list of active workers at ISA (both employees and subcontracted) that are registered with INSS. Paper records are available for the period prior to October 2008, but identifying INSS numbers in this fashion is very time consuming.

There are four sets of files at the ISA Hospital. "Active" files are maintained for individuals who are currently working at the Ingenio. "Inactive" files are maintained for workers no longer employed at the Ingenio, but it is possible that workers may have a record in both areas due to changes in work status. "Deceased" files are kept for workers who have passed away and "11-42" files are kept for workers during their first eight weeks of employment before their INSS number is issued. The latter applies to all workers *except* cane cutters, who are eligible for care from the moment of hire, and so they have an "Active" file upon employment.

B. Centro de Salud Chichigalpa Medical Records:
Records at the Chichigalpa Health Center are organized by a medical record number unique to the health center. There are two record categories at the health center including records of active patients and death records**.** We believe that inactive medical records are discarded after five years. In 2007, a nephrology specialist began offering services at the health center and compiling a database of CKD patients. In its current form, the database contains information on approximately 2,200 subjects and includes demographic information, medical record number and CKD status coded from 0 to 4, with 0 indicating that the patient recently transferred to the facility and 1-4 representing the stage of CKD.

C. HEODRA Hospital:
Like the Health Center in Chichigalpa, records at the HEODRA hospital in León are organized by a medical record number unique to that facility. The HEODRA hospital has computerized some medical record information of CKD patients, allowing a search of patients by name or other demographic information. However, not all CKD patients are included in this computer database, and the database is only accessible by year, not across years. These records are also stored in hard copy, listed by visit for the past 2-3 years. HEODRA also has record books of hospital discharges from as far back as 2005. Discharges are recorded by date and can be reviewed manually. HEODRA hospital does not keep inactive records after 5 years.

D. Other linkage sites:
Medical record keeping was also investigated to a limited extent at the Posoltega and El Viejo Health Centers. The Posoltega Health Center keeps a list of CKD patients that is electronic, but we were not able to review this database. Additionally, Posoltega keeps alphabetized cards of all patients. Cards indicate that a patient visited the health center, a medical record was established, and a medical record number was assigned. Medical record numbers from these cards can be used to search for full medical records. However, inactive medical records are discarded after 5 years. Additionally, Posoltega also keeps yearly chronological lists of deceased patients.

The El Viejo Health Center uses a similar recordkeeping method as the Posoltega Health Center. Patient cards are organized alphabetically and medical records are discarded after 5 years of inactivity. Additionally, the Hospital España sends lists of deceased individuals that indicate El Viejo as their primary health center. These lists can be reviewed manually to determine use of Hospital España for secondary health care.

V. Subject Selection for Pilot Study

Following a thorough review of the available employment and payroll records, we selected subjects for the Pilot Cohort Study. An individual was eligible for the study if employed at ISA sometime between July 1, 1997 and June 30, 2010, as either an ISA employee or contract worker in one of the following positions: cane cutter, cane gatherer ("pucho"), pesticide applicator, field machine operator, irrigator, or factory worker. ISA payroll records were the primary source for subject selection, as they were considered the best source for identifying workers in each job category over time. A total of 243 subjects were selected across job categories and time, including 90 cane cutters and 153 workers from the other occupations. A larger portion of cane cutters was selected because they are likely to be the main focus of a future cohort study. The remaining subjects were divided approximately evenly among the remaining five occupational groups. Table 2 describes the key features of the selection process. Payroll records from a specific month were used to select subjects employed at the midpoint of the annual work period (e.g., zafra) for each job category. Selection procedures for each job category across time are described in more detail in the Appendix.

Table 2: Subject Selection for Pilot Study

Type of Worker	Record Source for Subject Selection	Planilla[+]	Month of Selection	No. Targeted for Selection From Each Calendar Year (1997-2010)	Total No. Selected
Cane Cutter	Payroll, Enganche screening database and Ordenes de Pegue*, Harvest Office database*	10, 11	February	7	90
Irrigator	Payroll	5	January	2-3	32
Pesticide applicator	Payroll	5	June	2-3, except for 1997-1999**	26
Field Machine Operator***	Payroll	4	January	2-3	32
Factory Worker	Payroll	12, 13	February	2-3	32
Cane Gatherer	Payroll	4	February	2-3	31

[+] Payroll division used for processing workers' wages
* These databases include contract workers
** Records during these years appeared incomplete.
*** Includes mechanics, drivers, welders, carpenters

VI. Data Collection and Analysis for Pilot Study

A. Abstraction of ISA Employment Records: Once a subject was selected, his ISA ficha number was given to the employment office and his paper employment record from the start of employment through 2005 was abstracted manually, following a protocol for abstraction of

paper employment records. Employment records from 2006 through 2010 are electronic and so these records were abstracted by computer.

B. Abstraction of Payroll Records: Because of limited time and resources, a subset of 75 workers was selected for full abstraction of payroll records. A systematic selection across calendar years and ficha numbers was used to select these individuals. Paper payroll records from the start of employment through 1999 were abstracted by hand following an abstraction protocol and payroll records from 2000 through 2010 were abstracted electronically from available computer files.

C. Employment Data Collection for Cane Cutters: Cane cutters experienced an occupational transition between 2000 and 2003, which makes establishing their occupational history more challenging. During this period, cane cutters transitioned from all being direct hires of ISA to all being contract hires, their work negotiated through a third party contractor.

Like all study subjects, cane cutters were selected from the period 1997 to 2010. Because of variations in methods of employment and the available records, selecting and abstracting cane cutter employment history merits separate description.

There are a variety of data sources indicating the employment history of cane cutters as ISA employees and as contracted workers. Employment histories of cane cutters can be divided up into 3 periods (after 1997):

1. 1997-1998 zafra: All cane cutters were directly hired by ISA and therefore employment history throughout the zafra would be retrieved via ISA employment and payroll sources, as noted above.
2. 1998-1999 zafra through 2002-2003 zafra:
    a. A percentage of cane cutters were directly hired by ISA. Their employment history would be retrieved via the sources noted above.
    b. Some cane cutters were hired via contractors. Ficha or carnet numbers for these workers appear in the Ordenes de Pegue starting in 2000.
        i. In 2000 and 2001, employment can be assessed in the Ordenes de Pegue using ISA ficha number. Frequency of appearance in a year's worth of records is understood to represent days worked during a particular zafra.
        ii. From 2002 on, carnet is used. Since this carnet will be unknown if a worker was selected for the study during a year when they were an employee of ISA, their carnet can be obtained by matching information on cedula, INSS, or name from the ISA employee files to records in the enganche databases. Once a carnet is obtained, number of days worked per year can be assessed in the Ordenes de Pegue in the same manner as in 2000 and 2001.
3. 2003-2004 zafra through 2009-2010 zafra: All cane cutters were hired by contract.
    a. Employment can be assessed in the Ordenes de Pegue as described above for cane cutters selected as ISA workers.
    b. From 2003-2006, potential cane cutter subjects were identified using the Enganche Databases, and employment was then confirmed by searching for carnet in the corresponding year in the Ordenes de Pegue. Previous direct employment at ISA was determined by matching INSS, cedula and/or name in the ISA electronic database and obtaining a previously used ISA ficha. This ficha was then also used to evaluate paper records from before 2000.

c. Additionally, starting in 2007-2008 more detailed records of contracted cane cutter employment become available. Cane cutters were selected into the study directly from these records for the 2007-2009 zafras, contractor employment was assessed in other years using the Ordenes de Pegue, and ISA employment was assessed by the database matching method described above. The detailed cane cutter records could also be used to gather information on intensity of work, represented by average tonnage cut per pegue per day. Currently, this information is available to us as PDF printouts of work done each week for each worker in the 3 harvest seasons, organized by contractor group. In its current state, accessing information about specific workers given only their contractor ID would be unpractical, as a worker's hiring contractor may change from week to week and across zafras. However, it may be possible to access this information in database format.

D. Linkage of Employment and Medical Records: Once a subject's employment file was abstracted, a medical record linkage form was filled out to document linkage with medical records at the ISA hospital and other facilities. The linkage form includes name, date of birth and other identifying information such as the INSS number to enable correct identification of subjects in patient files at the linkage institutions. Linkage was attempted at the following institutions: ISA Hospital, Centro de Salud in Chichigalpa, HEODRA Hospital, and a few others.

E. Medical Record Abstraction: All available medical records were abstracted for the subset of 75 workers whose payroll records were also fully abstracted. Medical abstractors received training prior to medical record abstraction. They were instructed to abstract the entire medical record. Particular emphasis was placed on recording all visits relating to an infectious process; urinary tract infections; any type of renal disease; urinary symptoms; visits for which pain killers, NSAID's, antibiotics, or any other potentially nephrotoxic drug was prescribed; systemic illnesses (such as malaria, leptospirosis or dengue); pesticide poisoning; dehydration; and heat stroke. Additionally, special emphasis was placed on accurately recording vital signs and laboratory results related to the kidney and urinary tract (serum creatinine, urinalysis results, renal ultrasound results, etc).

F. Data Entry, Cleaning and Analysis: Several Microsoft Access databases were built to enable computerization of all abstracted data. These databases included features to reduce the occurrence of data entry errors. In addition, 100% verification, and various data cleaning procedures were performed before analyses were conducted. Relevant data from electronic records were also extracted and merged, as appropriate, with the Access files of the manually abstracted records. Merging procedures identified problems with overlapping records and inconsistent methods for documenting calendar time. All Access data files were converted to SAS (Statistical Analysis System) files to enable statistical analysis.

VII. Informational Interviews

Ten subjects (five past workers and five current workers) from the total of 243 selected for the Pilot Study were selected for in-person interviews. The interviews, which were conducted by Damaris Lopez de Pilarte, Study Coordinator, and Dra. Marta Pastora Arostegui, obtained information about the worker's employment history at NSEL and elsewhere, sources of medical care, medical conditions and treatments, and lifestyle characteristics. The goals of the interviews were to: (1) compare self-reported information about employment with the data obtained from record review, (2) identify sources of medical care outside ISA, and (3) explore collection of information about potential risk factors of CKD outside ISA.

VIII. Results of Pilot Study

Records for a total of 243 workers were examined for the pilot study.

A. Demographic Characteristics of Subjects

As expected, all subjects were males who were born in Nicaragua. Birthplaces were primarily Chichigalpa (63.2%), Leon (6.5%), Chinandega (5.0%), Posoltega (4.0%), and El Sauce (3.0%). The remaining 19% of subjects were born in 21 other municipalities located mainly in the country's western region. Addresses recorded in the ISA medical records indicate that the population is fairly stable: 85% resided in Chichigalpa at their most recent encounter.

As shown in Table 3, subjects were, on average, 21.4 years old upon hire at ISA. The average age at hire was fairly consistent across selected work groups, except for factory workers who were nearly one year older than other workers. While age at hire for most subjects was clustered around the average, the range was quite wide, spanning from the early teens to the late fifties. The average age of subjects is currently 38.2 years.

Overall, 64.0% of subjects had a body mass index (BMI) in the normal range (18.5-24.9 kg/m^2), indicating an appropriate weight for height according to commonly used definitions. However, the proportion of subjects with normal body mass index varied across job categories: Higher proportions of cane cutters (79.0%), cane gatherers (65.5%), pesticide applicators (69.2%), and irrigators (80.7%) had a normal body mass index as compared to field machine operators (32.1%) and factory workers (41.4%). In fact, the proportions of overweight (BMI 25.0-29.9 kg/m^2) and obese (BMI>=30.0 kg/m^2) workers were 46.4% and 17.9%, respectively, among field machine operators and 37.9% and 17.2%, respectively, among factory workers. Only 4.5% of the population was underweight (BMI<18.5).

B. Employment Characteristics of Subjects

As shown in Table 4, more subjects began their employment at ISA in the later years of the study period; the majority of subjects began their employment between 1995 and 2010. The termination of employment occurred mainly from 2000-2009. Slightly more than half of subjects (54.6%) are currently employed at ISA.

Most study subjects held a variety of jobs during their employment at ISA. For example, 48.3% of cane cutters held other positions at some point during their employment, including jobs as pesticide applicators (30.3%), irrigators (14.6%), cane gatherers (5.6%), and factory workers (3.4%). The duration of employment in these other jobs was generally shorter than the duration in the selected job category.

Overall, the total duration of employment at ISA was relatively short: 58.1% of subjects worked for 5 years or less and mean duration of employment was 5.8 years. However, total employment duration varied across job categories: Cane cutters, cane gatherers, and pesticide applicators had shorter average durations of employment (3.7, 3.8, and 5.5 years, respectively) than field machine operators, irrigators and factory workers (7.4, 7.2 and 10.8 years, respectively).

Like total duration of employment, employment duration in the selected job category also varied by job. Cane cutters, cane gatherers, and pesticide applicators had shorter average durations of employment in their selected job category (2.7, 1.8, and 2.9 years, respectively) than did field

machine operators, irrigators and factory workers (6.0, 4.8 and 8.5 years, respectively). Overall, subjects worked about half of a calendar year at ISA (mean: 6.1 months). Cane cutters worked for a slightly fewer months of the year (mean: 4.5 months) than other workers (5.1-7.6 months).

C. Linkage Rates for Employment and Medical Records

As previously described, medical records from the ISA Hospital, the Centro de Salud in Chichigalpa, Hospital Escuela Oscar Danilo Rosales Arguello (HEODRA), and a few other medical facilities were linked to employment records.

At the ISA Hospital, the INSS number represents the medical record number. Table 5 describes the linkage results by job category for ISA Hospital records. The results for cane cutters are separated by calendar year because of their transition from ISA employees to contract workers. ISA linkage rates were very good for all job categories (94-100%), except for contracted cane cutters (46-57%) during the period 2003 through 2010.

A number of factors might contribute to the low linkage rate among these contracted cane cutters. Overall, cane cutters worked at ISA for shorter periods of time as compared to other workers. Therefore, it is possible that they make less use of the healthcare system than workers with other jobs due to their shorter duration of work. Other possible explanations include a higher prevalence of lost, misfiled, and discarded medical records, and unavailable INSS numbers (needed for linkage) among these workers.

Linkage rates were also low at all other medical facilities for both ISA employees and contracted workers. Overall, the linkage rates were 13.6% at Centro de Salud in Chichigalpa, 0.4% at HEODRA, and 0% at other facilities. Possible reasons for these low linkage rates include the organization of the medical record files by a number unique to the health care facility (with no cross-indexing by patient name) and the practice of discarding records after five years of inactivity.

D. Medical Characteristics of Subjects

According to the Study protocol, we attempted to abstract the ISA Hospital records of 75 subjects. We were able to locate the records of 61 of these subjects (81%) for abstraction. Medical encounters were dated and signed, and usually contained the worker's vital signs, a reason for the visit, diagnoses, medications prescribed and laboratory results when indicated. Information that was often missing from the record included past medical history, social history, family history of CKD and other diseases, and drug allergies.

Time required for abstraction was highly variable and dependent on the length of the medical record. On average, it took approximately one hour to fully abstract a single record. The major difficulty encountered during medical record abstraction was understanding the handwriting in some of the records.

1. Healthcare Utilization

Table 6 describes the average annual number of medical encounters by job category. Overall, the average number of encounters ranged from 1.0 to 5.2 per year, with a median of 1.7. Irrigators and pesticide applicators had more encounters (median=2.0) than factory workers and cane gatherers (median=1.4-1.5). Table 7 describes the number of medical encounters from 1965 through 2010. Utilization was fairly stable over this period (median 2.0-3.0), except there was a period of increased utilization during 2000-2009 (median: 6.0 encounters).

Workers began to utilize the healthcare system, on average, 6.3 months after starting work at ISA; however, there was a wide range in the timing of first utilization (from days to several years). Seventeen workers (27.9%) had a medical encounter before their first employment date, which was likely to be a pre-employment general medical examination.

2. Medical Conditions

As shown in Table 8, medical conditions diagnosed among workers were quite varied. Overall, the mandatory pre-employment medical examination represented 21.8% of all encounters. Another 21.0% of encounters were due to infectious diseases (not including urinary tract infections), making this the most common condition for which workers sought medical attention. Among the infectious diseases, upper respiratory tract infections comprised 13.8% of encounters and febrile syndrome comprised 2.8% of encounters. Disorders of the kidney and urinary tract accounted for 15.6% of medical encounters, with 80% of these diagnosed as urinary tract infections. Disorders of the joints and adjacent tissues accounted for 9.1% of medical encounters. It should be recognized that medical diagnoses were not usually supported by laboratory or radiologic evidence.

There were 26 workers (42.6%) with an encounter related to an occupational trauma or burn. Five workers (8.2%) had a diagnosis of hypertension, and two (3.3%) had a diagnosis of diabetes. Eleven workers (18.0%) had at least one diagnosis of renal failure not otherwise specified, and two workers (3.3%) had at least one diagnosis of acute renal failure. Three workers (4.9%) had at least one diagnosis of poisoning from heavy metals or pesticides.

3. Urinary Tract Infection

Table 9 describes the frequency of urinary tract infection (UTI) diagnoses among subjects. Overall, 42 of the 61 workers (68.9%) were diagnosed with a UTI on at least one occasion. The diagnosis of UTI was recurrent, with workers receiving a diagnosis of UTI up to 17 times (median=1.5). There were a total of 140 medical encounters resulting in a UTI diagnosis. Dysuria (with or without fever and back pain) was present in 40.7% of medical encounters with a diagnosis of UTI, back pain (with or without fever) in 21.4% and fever in 5.7%. Of note, no mention of dysuria, fever or back pain was recorded in 32% of the encounters that resulted in a UTI diagnosis. Laboratory evidence for UTI includes urine dipstick findings of nitrite and leukocyte esterase positivity, pyuria on microscopic examination of the sediment and/or a positive urine culture. Only 61 of the 140 medical encounters (43.6%) where a diagnosis of UTI was given had a urinalysis accompanying this diagnosis. This low proportion of encounters with a urinalysis might stem from UTI diagnoses being made based solely on clinical findings, or insufficient documentation in the medical record. Among the 61 encounters with a UTI diagnosis and a urinalysis, only 33 (54.1%) had evidence of WBCs in urine microscopy, and only 1 (1.6%) had evidence of positive nitrites in urine dipstick. Leukocyte casts were present in 11 (18%) of the 61 encounters. Table 10 provides a detailed description of urinalysis results among workers with abstracted medical records.

4. Chronic Kidney Disease

We define CKD in this report as one of the following: (a) two serum creatinine levels of >1.4 mg/dL at least three months apart, and (b) CKD stated as a diagnosis in the medical record. Table 11 provides a comparison of CKD diagnoses based on the two definitions. Inconsistencies between definitions in ascertainment of CKD cases occurred in two occasions: A case by definition (a) that was noted in the medical record as renal failure not otherwise specified and a case by definition (b) that did not satisfy criteria for definition (a).

Of note, there was another case of CKD who satisfied our criteria for definition (a) since he had two serum creatinine levels of >1.4 mg/dL at least three months apart. However, the creatinine values decreased below 1.5 mg/dL afterwards. Therefore this case may be misclassified as having CKD. This same case had also been ascertained as having CKD according to definition (b).

Seven subjects (11.5%) were identified as having CKD. Using the creatinine-based definition, subjects with CKD had a median age at diagnosis of 34 years and a median work duration at ISA of 14.2 years prior to their diagnosis. Table 12 provides information on CKD prevalence according to the longest job held at ISA, and Table 13 provides information on CKD prevalence according to the selected job. Twenty-nine percent of workers whose longest job was cane cutting had a diagnosis of CKD. Lower CKD rates (0-10%) were observed among workers whose longest job was in other categories.

As noted previously, only 8.2% and 3.3% of subjects had a diagnosis of hypertension and diabetes, respectively. Among the seven subjects with the creatinine-based definition of CKD, one worker (14.3%) had a diagnosis of hypertension and none had a diagnosis of diabetes before CKD was diagnosed. Table 14 provides the mean systolic and diastolic blood pressure for all workers and those who developed CKD (before their diagnosis was made).

Proteinuria has been identified as a risk factor for progressive kidney disease. Among workers that developed CKD, 3 of 7 (42.9%) had proteinuria determined by urine dipstick before they were diagnosed with CKD (2 of them had a protein concentration of 30 mg/dL and one of them had a concentration of 100 mg/dL). It is uncertain whether these proteinuria determinations occurred during symptomatic UTI's, which may have resulted in transient proteinuria. Table 10 provides a detailed description of protein concentration determined by urine dipstick among the 61 subjects regardless of CKD status.

CKD encompasses a spectrum of different pathophysiologic processes associated with abnormal kidney function. A glomerular process is characterized by large amounts of protein in the urine, in contrast to a tubulointerstitial process characterized by white blood cells, leukocyte casts, and lesser amounts of protein in the urine. Among the seven workers who developed CKD, urinalysis results before diagnosis showed at least one episode of white blood cells in 57.1%, and at least one episode of leukocyte casts or proteinuria in 42.9%.

5. Medication Prescriptions

Table 15 provides a detailed listing of the most commonly prescribed medications. Antibiotics and non-steroidal anti-inflammatory drugs (NSAIDS's) are the most commonly prescribed class of medications, accounting for 33% and 25% of prescriptions, respectively. Among the medications more commonly associated with nephrotoxicity, NSAID's were prescribed to workers up to 3.5 times per year (median=1.0). It is important to note that the use of NSAID's is likely underestimated among workers, since these medications can be obtained over the counter. The aminoglycoside antibiotics were prescribed to workers up to 1.3 times per year (median<1.0).

6. Summary of Medical Record Abstraction

The main findings of the medical records abstraction are as follows:

1. In men, published epidemiologic data indicate that the prevalence of UTI's is less than 1% until the sixth decade of life (Shaeffer 1994). Therefore, the high number of UTI diagnoses among workers is unusual. It seems that a high number of UTI's are

diagnosed solely on clinical findings since only 43.6% of medical encounters with the diagnosis of UTI had a urinalysis. Since urine cultures were not performed in the vast majority of workers with UTI, confirming the diagnosis is extremely difficult. We believe that UTI is most likely over-diagnosed among workers; however, this warrants further investigation.

2. A total of seven workers (11.5%) were identified as having CKD. These cases had a median age of diagnosis of 34 years following 14.2 year median duration of work at ISA. The frequency of CKD was greater among subjects whose longest job was in cutting cane. The high proportion of workers with CKD who had evidence of pyuria, leukocyte casts, and low grade proteinuria are suggestive of a tubulointerstitial process rather than a glomerular process. However, due to the small number of workers with CKD, these findings should be confirmed in a larger sample.
3. The prevalence of diabetes and hypertension was relatively low.
4. Antibiotics were the most frequently prescribed medication, and NSAIDs were the next most common.

Because of the small number of medical records abstracted, caution is warranted in extrapolating these findings to all ISA workers.

E. Results of Informational Interviews

As previously described, ten subjects from the total of 243 selected for the Pilot Study were selected for in-person interviews. Five of these interviews were conducted among current workers and five were conducted among former workers. The selection of these subjects was not random and so the following results should not be considered representative of all current and former ISA workers.

1. Demographic and Lifestyle Characteristics

As described in Table 16, interviewed subjects ranged in age from 28-70 years. As expected, former workers were older and more likely to be married than current workers. The consumption of tap and bottled water and alcoholic beverages; cigarette smoking, and raising crops for food were fairly common (50-90%) in both groups. Former workers were more likely than current workers to consume well water (60% vs. 0%) and raise animals for food (100% vs. 20%) while current workers were more likely than former workers to drink lija (60% vs 20%) and use pesticides on home grown crops (40% vs 20%).

2. Work Characteristics

As described in Table 17, many current and former workers held a variety of jobs at ISA including work as cane cutters, herbicide and pesticide applicators, machine operators, irrigators, cane gatherers, and factory workers. Many interviewed workers also reported that they, at some point, held jobs outside of ISA, including jobs as drivers, brick layers, farmers, security guards and pest controllers. Former workers were more likely than current workers to have held more than one type of job at ISA. Self-reported exposure to pesticides, herbicides, and dust while working at ISA or elsewhere were commonly reported (70-90%) in both groups. In contrast, self-reported exposure to solvents was rare (10%) and no exposure to lead, cadmium, or gasoline were reported.

When we compared self-reported job histories with those found in ISA employment records, we found a moderate level of agreement. Generally, two to four jobs reported by a subject were also noted in the employment records. However, self-reports often failed to mention some jobs

while employment records failed to mention others. For example, one of the current workers stated that he held jobs as a cane cutter, pesticide applicator, machine operator and field supervisor. In contrast, his employment record only noted jobs as a cane cutter, pesticide applicator, and weeder.

3. Medical Characteristics
As described in Table 18, numerous medical conditions were commonly reported by interviewees, including a history of chronic renal insufficiency (40%), elevated creatinine levels (60%), kidney or bladder infection (70%), chistata (70%), dehydration (40%), hypertension (40%), and heat stroke (30%). A family history of CKD was also common in both groups (60%) while a personal history of diabetes was uncommon (10%). Use of the antibiotic gentamicin and folk medicine were fairly common (20-30%) and use of NSAIDS was widespread (90%).

4. Sources of Medical Care
As shown in Table 19, all 10 interviewees reported using the ISA Hospital for medical care. Other commonly used medical facilities were Centro de Salud in Chichigalpa (60%), Hospital Escuela Oscar Danilo Rosales in León (HEODRA) (40%), and Hospital España in Chinandega (30%). Current workers also reported using several other sources of medical care, including Hospital Mauricio Abdalah in Chinandega, Amorsa in Chichigalpa, a private doctor in Leon, and "la consulto" in Chinandega. None of the former workers reported these other sources.

IX. Recommendations and Conclusions

Based on the information obtained in the pilot study, we have concluded that a retrospective cohort study could feasibly be undertaken and that such a study could represent a valuable contribution to an understanding of the causes of CKD. At the same time, despite the strength of much of the ISA record system, limitations in the availability, completeness, and quality of data have forced us to modify our original plan and—of more importance to the Dialogue Table participants—reassess our expectation of the ability of the study to answer the question of a possible causal connection between work at ISA and the occurrence of CKD with the speed and level of confidence desired by the parties in the Dialogue Table. We have reached the conclusion that a retrospective cohort study would likely produce results that would take too long to generate and would leave too much uncertainty to satisfy the needs of the parties in the Dialogue process.

Our perspective is also influenced by the information obtained through our walkthrough and assessment of the work practices at ISA and review of the medical literature on the state of knowledge of the relationship of pesticide exposure and CKD. This information was presented in the August 2010 report "Industrial Hygiene/Occupational Health Assessment: Evaluating Potential Hazards Associated with Chemicals and Work Practices at the Ingenio San Antonio (Chichigalpa, Nicaragua)". The report found that of the five possible occupational causes of CKD that we considered, only one of them (heavy metals) has ever been shown to cause CKD, and it appears that widespread exposure to heavy metals is unlikely. Any of the remaining agents could cause CKD, but the task of demonstrating a causal relationship is more challenging than it would have been had a causal connection already been proven in another setting.

Although a single study probably would not in and of itself sufficiently resolve the specific questions posed by the Dialogue participants, a retrospective cohort study would still provide value to the overall research process. The records at ISA and nearby health centers represent a

unique resource, and their analysis would make an important contribution to scientific knowledge and improve our understanding of the causes of CKD in Nicaragua.

This section presents the reasons for our conclusions and the basic design of a proposed retrospective cohort study.

Data strengths and limitations and impact on proposed study design

Below is a description of the main factors that we considered in proposing a retrospective cohort study design at ISA. Ideally, available data would support the following steps in the research process:

1. Identify the records of a group of workers who worked at ISA during a particular time period and categorize them according to agents or conditions to which they might have been exposed, which were enumerated in the Industrial Hygiene Report of August 2010 (pesticides, heavy metals, volume depletion/muscle damage, systemic infections, and silica). In order to be able to determine whether these agents or conditions are associated with CKD, there must be a sufficient difference between groups of workers in exposure to these factors.
2. Locate medical information on these workers that would allow us to determine whether or not a case of CKD had occurred.
3. Link the exposure and disease records, and calculate the rate of occurrence of new cases (incidence rate) of CKD for each exposure group. The records should have sufficient information to establish that exposure occurred before disease.
4. Compare the incidence rates for different exposure groups to get a measure of whether different levels of exposure were associated with an increased rate of CKD.
5. Collect information on other causes of CKD (confounders) that could differ in frequency by exposure group, which would be used to adjust the rate comparisons so that they would be more accurate.
6. Ascertain enough cases of CKD in each exposure group so that the estimates derived from the rate comparisons between different groups would be relatively stable.

Data strengths and limitations almost always influence the design and feasibility of epidemiological studies. Therefore, based on the information obtained in the pilot study, we had to identify strengths and limitations of a retrospective cohort study, determine what could be done to maximize the strengths and minimize the limitations, and then assess the value of conducting the study, given the limitations that would inevitably remain.

The main strengths of the information available at ISA include:
1. Availability, completeness, and quality of employment, payroll, and medical records over substantial periods of time for ISA employees.
2. Availability of electronic employment and payroll records since 2006 and 2000, respectively.
3. Mandatory pre-zafra screening for serum creatinine of applicants for temporary employee positions since 2000 and for applicants for contract cane cutter positions since 2003.
4. Records of past workers with CKD, which, although not complete, would facilitate the identification of workers who developed the disease.

The main limitations identified were:
1. No information on contracted workers other than cane cutters.
2. Variation over time in availability, completeness, and quality of information on contracted cane cutters.

3. Differences in utilization of medical care at ISA based on contractor status.
4. Difficulties in finding medical records at external health care facilities for ISA employees, due to use of facility-specific identification systems and the practice of discarding medical records after five years for deceased or inactive patients.

Based on consideration of these factors, the proposed cohort study design would include the following criteria:

1. Contracted workers other than cane cutters would not be included in the study because records are too limited for these workers.
2. The main job categories to be included in the study would be cane cutters, pesticide applicators, irrigators, machine operators, "puchos", and factory workers.
3. We would need to institute improved and expanded methods to identify CKD among those workers chosen for the study.

Proposed design of retrospective cohort study

The proposed study would include approximately 5,000 persons who have worked at ISA using the following basic study design:

1. Identify and link the employment and payroll records of workers in selected job titles at ISA. With these linked records, we would construct work histories on ISA workers that include type of job, number of years worked at ISA, proportion of year worked at ISA, average number of hours worked per week, and amount of tonnage cut.
2. Classify different jobs according to their potential for exposure to agents and conditions discussed in the 2010 Industrial Hygiene Report,
3. Identify which of these workers developed CKD, based on information in medical records at ISA and external medical facilities. Identification of workers' medical records in the latter location would require creation of electronic databases and active community ascertainment efforts.
4. In order to control for the possibility that exposure to agents or conditions outside of ISA are actually responsible for differences between groups of workers (confounding), conduct detailed abstraction of ISA and other medical facility medical records, and administer questionnaires, on a subset of former and current workers.
5. Based on all the above information, compare the rate of CKD in different groups of workers and assess whether there are patterns that are consistent with different exposures.

Workers selected for the study would meet the following eligibility criteria:

1. Worked at ISA for more than one zafra year. (Purpose: exclude workers who worked only a very short time.)
2. Worked at ISA in the 2004-2005 zafra year or later. (Purpose: begin recently enough to improve chances of determining whether a worker developed CKD while working or after leaving work at ISA. Begin early enough to capture enough cases of CKD to permit a stable analysis.)
3. Began working at ISA in 1985 or later. (Purpose: prior to 1985, job titles in employment records are not specific enough to distinguish different types of jobs.)
4. Free of CKD as of 2004-2005 zafra year or date of hire, whichever occurred later. (Purpose: study should only include people who did not already have CKD at the beginning of the study.)

Major study components include: *exposure, outcome, and confounders*. The following sections present descriptions of issues identified during the pilot study and plans for remedying these data gaps.

Exposure
We lack direct information on exposure to the agents and conditions of interest as identified in the Industrial Hygiene study report of August 2010 (pesticides, heavy metals, volume depletion, systemic infections, and silica), so we would instead need to use job title as a proxy measure. This is not unusual for a retrospective occupational cohort study, and is reasonable as long as the level of exposure significantly varies by job category. Based on our understanding of the jobs at ISA and individuals' work histories obtained through the pilot study, it should be possible to classify jobs according to whether workers have higher or lower average exposure to the different agents and conditions.

To link the six job categories to the agents and conditions of interest, we would use a job-exposure matrix, which involves creating a table for each exposure of interest with rows and columns indicating the job title and calendar year, respectively. The value in each cell represents the level of exposure to each agent, which is allowed to vary over calendar time to reflect changes in such factors as usage, work practices, and personal protective equipment. If the intensive worker observation study proposed in the Scoping Study were to occur prior to analysis of the data, the study results would also inform the rating of job titles, particularly in relation to agrichemicals, volume depletion, heavy metals, and silica. Otherwise, knowledge about the exposures over time would be derived from available records, knowledge about exposure occurring on jobs, and input from industrial hygienists, company supervisors, and current and former workers.

Individual work histories would be determined for all workers in the study. From payroll and employment files for ISA employees, available information includes total years worked at ISA, years worked in different job categories, calendar years worked, total number of days and hours worked per zafra, and hours worked per day and days worked per week. For cane cutters, a combination of employee payroll records, enganche testing data, and ordenes de pegue and related contractor payroll records would be used to construct the same work histories as for the employees, with the exception of hours worked per zafra or per day. In addition, for years available, which we currently understand to be 2007 and later, we would categorize cane cutters by amount of tonnage cut.

Based on these work histories, potential occupational exposure to the agents of interest would be determined for individual workers by linking their work histories to the job-exposure matrix and creating summary measures for the agents, such as duration of exposure and cumulative exposure to the agent.

Outcome
The primary definition of whether a participant developed CKD would be based on: (1) the presence of a CKD diagnosis in the medical records of the ISA Hospital or an external health facility such as the Centro de Salud in Chichigalpa, (2) two estimated glomerular filtration rate (eGFR) measurements of <60 ml/min at least three months apart, or (3) two creatinine measurements of ≥1.5 mg/dL at least three months apart.

We would not be able to determine whether the great majority of workers in the study have developed CKD based solely on the records available at ISA. This situation is also not unusual,

but it could result in a biased estimate of the association observed between exposure and CKD. There are at least three scenarios that might result in insufficient information:

1. Workers may simply stop working at ISA for a variety of reasons unrelated to CKD, and most of these workers will not have developed CKD at the time they stopped working. However, it is possible that they may have developed CKD after leaving employment.

2. Workers could have elevated creatinine levels at the pre-employment screening, and therefore not be hired for that zafra. Although a single elevated creatinine level is indicative of likely CKD, it does not mean that a worker necessarily has CKD and therefore does not meet the case definition.

3. Workers may also have a high creatinine test while they are working. For cane cutters and pesticide applicators, there is a voluntary testing program during the zafra, and most workers participate. Other workers would only receive a creatinine test if they go to the ISA Hospital for any reason. Since CKD is generally asymptomatic until later stages, it is unlikely to generate physical complaints that would lead to a medical visit. Therefore, cane cutters and pesticide applicators have a greater opportunity to have an elevated creatinine test. Follow-up after an elevated creatinine test also differs by job category. A temporary employee with a single elevated creatinine would have additional testing to establish a diagnosis of CKD at ISA, while any additional testing for a cane cutter would occur outside ISA and be unavailable in the company medical records. Therefore, given a single elevated creatinine level during the zafra, we are more likely to have a CKD diagnosis for temporary employees than for cane cutters.

It is therefore necessary to follow these workers outside ISA to determine whether they developed CKD. We were hoping that we could identify most of them at a relatively small number of medical facilities. However, we found that the recordkeeping systems at these facilities were not compatible with the information that we would have on each worker from their ISA records. We therefore propose to take the following steps to capture the majority of workers:

1. Create an electronic database of all existing medical records at a small number of facilities. Chichigalpa Health Center would definitely be included. We would determine the other facilities by conducting a survey of former and current workers regarding where they seek their medical care. By searching the database, we would be able to locate the medical records of many of the workers. Note that development of these files would aid not only this study but also future exploration of hypotheses among users of these facilities who haven't worked at ISA.

2. Among those who are not identified through the electronic databases, we would visit their last known address based on ISA records. If the worker or his/her family still lives there, we would ask for the information needed to locate their medical record. We could also consider a creatinine test among those who consented. If the family no longer lives there, we would ask neighbors about their current residence and attempt to locate them. The information obtained would not only permit us to directly ascertain more outcomes but would also provide data for imputation of outcome among individuals who were unable to be contacted.

Confounding

There are a number of factors unrelated to work at ISA that have been suggested as possible causes of CKD in Western Nicaragua. If these factors are both fairly common and unevenly

distributed among workers in different job titles, it could bias estimates of the association between exposure and CKD. We can obtain information on a few of these factors from medical records (e.g., diabetes, pre-existing hypertension). We can also get relevant information from the biological testing study. For example, nonoccupational exposure to heavy metals could be a cause of CKD, but if we find very few workers with any elevation of heavy metals, it is less likely to play a role. Finally, we can review the questionnaires completed by cane cutters at the preenganche for information on occupational exposures in other settings, medical history, and other lifestyle behaviors. We would need to be cautious in our interpretation because the responses were given as part of the job application process.

The approaches described in the preceding paragraph would contribute but would not be sufficient to fully address concern about possible confounding. Therefore, we would also collect questionnaire data from study participants to get information on potential confounders. For example, one important question is what workers do and what types of exposures they have during the period of the year when they are not working at ISA. By collecting detailed data from workers across different job titles, we can assess the potential for confounding from this factor. In addition to obtaining information relevant for assessing and controlling confounding, the questionnaire would also serve as an opportunity to collect information necessary for the community-based outcome assessment approach described above for those participants interviewed.

Both the detailed review of medical records and questionnaires for all participants require many hours and therefore have a significant impact on the time and money necessary to complete a study. Therefore, we propose that medical records be reviewed for information on CKD occurrence for all 5,000 participants, but only a subset of study participants (e.g., 1500) have their medical records subject to detailed abstraction. Similarly, we would administer the questionnaires to elicit information on potential confounders from only a subset of study participants. The data obtained from these two subsets would be imputed to the entire cohort in the analysis. This design is known as “two-stage sampling”, and is meant to improve the efficiency of the study and reduce costs.

Data Analysis

After the information is collected on exposure, outcome, and confounders, we would analyze the data by comparing the rate of CKD in different groups of workers and see whether there are patterns that are consistent with different hypotheses. Given that there are likely to be workers for whom only a single high creatinine measurement is available even after our attempts at more intensive follow-up, we would also conduct a secondary analysis that includes individuals with a single high creatinine measurement but without the follow-up measurement necessary to meet the primary outcome definition for CKD.

The results would provide insight into whether there is any evidence that exposure to agents or conditions at ISA may be contributing to CKD among the workers there. Furthermore, the patterns may suggest that there are certain agents or conditions that are more likely to be the cause. For example, the groups at highest risk for CKD should differ based on whether volume depletion, infection (such as leptospirosis), or pesticides are more likely to be causes, although we can not count on the results to be sufficiently definitive to permit this detailed level of interpretation. However, regardless of the results, the wealth of information that would be collected and analyzed would represent an extremely valuable contribution to the research into the causes of CKD in Western Nicaragua.

Proposed study Cost and Timeline

We estimate that the proposed retrospective cohort study would take 2.5 years to complete, including analysis and report writing, and would cost approximately $1,000,000 USD. If we are able to collaborate with a Nicaraguan institution, some of the costs may be reduced.

Alternative Approaches to the Proposed Retrospective Cohort Study

A major reason for the cost and time required for the study is the cohort design, which requires approximately 5,000 people. The main alternative to the cohort design is the case-control study. Case-control studies are more efficient in terms of time and money than cohort studies because they generally require many fewer people than a cohort study. However, the efficiency of a case-control study generally requires two conditions: (1) most of the participants do not develop the disease (i.e., CKD), and (2) the people with CKD are easily identified or the information on CKD is relatively inexpensive to obtain compared to the exposure (i.e., job title) information. The first condition holds in this situation: we do not know the exact percentage of ISA workers in the job titles that would be included in the study who develop CKD, but it certainly seems to be below 20%. However, the second condition is not true in this situation. As described in the *Outcome* section, identification of people with CKD would be a laborious process, while the job title information is easily and inexpensively available through employment records. Furthermore, because only certain job titles would be included, it is more efficient to first restrict the study to people in those job titles. In the setting of this study, such an approach is possible with a cohort design, but not with a case-control design.

We also considered a modified case-control design in which the CKD cases were limited to those with records at the ISA Hospital. This would make the people with CKD easy to identify. However, because the likelihood of appearing in the hospital records is related to contractor status, we were concerned that the estimate of the relative distribution of job titles in the cases and controls would be biased.

A different approach would be to conduct a prospective cohort study. Instead of going backward in time and looking at cases of CKD that began occurring in 2004-2011, we would instead begin with current employees without CKD and follow them forward through time, beginning in 2012. The advantage of this approach is that we would have a stronger study that would provide us more confidence in being able to draw stronger conclusions from our results. First, we would be able to include more job titles because records at ISA have been getting better and we could also ask for further specific improvements. We can set up systems ahead of time so it would be easier to determine who has CKD and to obtain information on confounders. Finally, we would be able to conduct more detailed exposure measurements on current working conditions so we would have a better understanding of specific exposures. However, a prospective cohort study would likely take about 5 years to complete and cost perhaps 2.5 million dollars. Therefore it is unlikely to meet the needs of the Dialogue process.

Based on the results of biological sampling conducted by BU on workers in other industries (including miners, construction workers and port workers), we understand that the occurrence of the disease is common in other industries. Although the current proposal is restricted to ISA, it would be valuable to expand the study to other industries and/or additional ingenios. In addition to the increased cost, an expansion to new sites would require that a record system be available similar to that at ISA. It is doubtful that many other entities would meet this requirement, and a similar assessment as was done at ISA and described in this report would be necessary if there are indeed other candidate sites.

RECOMMENDATIONS

1. While we have attempted to make the study design as efficient as possible, we understand that the proposed study would require a major commitment of time and resources. Prior to making any final decisions, we recommend:
    a. Engaging one or two independent experts in occupational studies to review the proposed design.
    b. We have proposed two new methods which were not anticipated at the time we conducted the pilot cohort study and so are untested: (i) construction of electronic databases at Health Centers and community outreach to identify which participants have developed CKD and (ii) collection of questionnaire information from current and former workers to assess outcome and confounders. There should be pilot testing of these two methodologies in order to assess their feasibility and effectiveness. The information collected on CKD would also provide a better basis for assessing the number of cases likely to be identified over the period 2004-2011, which then would determine the number of workers who would need to be included in the study.

The following recommendations can be considered for implementation regardless of the decision to fund a retrospective cohort study:

2. The original Scoping Study proposed a large-scale detailed medical record review to assess the nature of the disease, its progression, and other potential causes. That study is implicitly rolled into the proposed retrospective cohort study, because approximately 1,500 medical records at ISA and external medical facilities would be getting reviewed already. The abstraction of medical records both at ISA and at community medical facilities in different areas of Western Nicaragua would be a useful task even if the complete retrospective cohort study is not implemented. More than 200 records at ISA will already have been abstracted as part of the pilot cohort. We would abstract perhaps 100 more records at ISA and approximately 400 additional records at 2-3 health centers in Western Nicaragua.

3. In the Industrial Hygiene Report of August 2010, we included the following recommendation regarding recordkeeping for contracted workers:

    > We recommend maintaining a recordkeeping system for contracted workers ("contratados"). Currently, there is almost no employment and limited medical information on this group of workers who constitute a substantial proportion of the workforce. Responsibility for employment recordkeeping appears to be left to the contractors ("contratistas"), but it is not clear what information is maintained and none of this information appears to be provided to ISA. As a result, it is difficult, if not impossible, to conduct basic surveillance of the health of all people working at ISA. ISA already has a good tracking system for its employees. It would be best if contracted workers could be brought into a single integrated system, but if this is not possible, an easily accessible parallel system would be the next best solution. It should be as simple to access the employment history of a contracted worker as it currently is for ISA employees.

    The importance of this recommendation was underscored by our experience with the pilot cohort study. We were unable to include any of the contract worker job titles other than cane cutters (and even the latter job title required piecing together different data sources that are themselves inadequate for identification and tracking). It is our understanding that certain steps have been taken in the last year to address this problem, but we don't know

the details, since this was not relevant for consideration of the retrospective cohort study (although it would be important information for consideration of a prospective cohort study). Because of the importance of being able to track whether the rate of new cases of CKD is increasing or decreasing among workers at ISA, we recommend that ISA obtain an independent analysis of their systems for collecting information on workers (both employees and subcontracted workers) and whether the data are sufficient for health surveillance.

4. Better linkage between the records at ISA Hospital and external medical facilities would have greatly facilitated follow-up of former employees and contracted workers. A modest investment would allow electronic record systems to be put in place at Health Centers in the area and allow better identification of CKD cases among all members of the community regardless of where they had been employed.

CONCLUSION

In the Scoping Study report prepared by the BUSPH team for the Dialogue Table in December 2009, one of the key activities proposed to explore potential causes of CKD was a retrospective cohort study of individuals who had worked at ISA. This recommendation was made based on a basic, but limited, understanding of the employment and medical record information and systems at ISA. Because the proposed study would be costly and was estimated to require 2½ years, a lengthy time period from the perspective of the Dialogue participants, a pilot study was undertaken in order to:

- Determine whether a retrospective cohort study is actually feasible;
- Assess the strength of evidence regarding any connection between work at ISA and CKD that would likely be gained from such a study;
- Outline the methods by which a retrospective cohort study could be conducted, along with estimated cost and duration; and
- Collect information from medical records that might inform hypotheses regarding the causes of CKD and inform future studies.

The pilot study was conducted September 2010 – May 2011 and involved extensive interviews with ISA personnel responsible for various record systems, abstraction and linkage of data from a sample of individuals who had worked at ISA, and analysis and interpretation of the data. We found that the record system at ISA had many strengths and had improved over time. In particular, we were able to construct employment histories for several categories of workers and were able to link employment records to medical records. There were two main limitations in the ISA record system from the perspective of conducting a retrospective cohort study: 1) lack of detail in job titles during the 1960s-1980s and 2) lack of any employment information for all categories of contracted workers, with the exception of cane cutters for whom the information was still less available and of lower quality than for employees. A third important limitation external to ISA was the inability to locate medical records at health centers or hospitals based on the identifying information obtained in ISA records.

After reviewing the data, we have concluded that it would be feasible to conduct a retrospective cohort study with the following restrictions: (1) only temporary employees and cane cutters could be included in the study, and (2) follow-up for CKD occurrence should begin with the 2004-2005 zafra. Furthermore, the conclusion was based on the likelihood that two additional activities would provide key data on possible confounders and the number and identity of individuals who developed CKD: construction of electronic databases at several health centers

and extensive community outreach. While it appears that these methods could be successful, they should themselves be tested before committing to a large-scale study.

We have also concluded that a retrospective cohort study would make an important contribution to determining the causes of CKD in western Nicaragua. Despite limitations in the ISA records, the amount of information that is available presents a unique opportunity to collect important data, which when combined with evidence from other studies would provide important scientific knowledge. The approach and methods of such a study are outlined in this report. A cohort study might be even more powerful if similar records were available at other sugar cane companies and/or companies in other industries, but that possibility has not been explored yet.

Although BU is confident of the broad value of a retrospective cohort study, our mandate from the Dialogue Table participants was to also assess the cost, duration, and strength of the evidence likely to emerge from a study, so that the partners could consider its potential value compared to other activities that the participants could undertake to address their concerns. The estimated duration of 2½ years has not changed from the original estimate in the Scoping Study. Based on the pilot study, the estimated costs are higher than originally estimated. Finally, we have lowered our assessment of how likely that a conclusion can be reached concerning an etiologic link between occupational exposure and CKD that would be of sufficient definitiveness for the Dialogue Table stakeholders. This reassessment is a result of both challenges discovered in the pilot study and other evidence obtained during the intervening time period which suggests it is possible that a cause not previously described may be responsible for the excess CKD. Our original assessment of the likely strength of evidence produced by a study did not guarantee a definitive opinion, nor does the current assessment rule one out. However, we have concluded that the likelihood of being able to offer a sufficiently strong opinion is less than previously believed.

In addition to providing information about feasibility and strength of evidence regarding a possible future study, the pilot study also provided an opportunity to gain new knowledge relevant to possible causes of CKD. Data from the medical record abstraction revealed a low prevalence of diabetes and hypertension, which are common causes of CKD in other settings. The data also added to the evidence from other sources that the underlying cause of the CKD represents a tubulointerstitial rather than glomerular process, based on the findings of low grade proteinuria and elevated white blood cells in the urine.

Finally, the medical record abstraction documented evidence of the frequent occurrence in younger men of dysuria (painful urination), also referred to as “chistata”, a condition that had been reported to our team informally by both patients and clinicians, as well as in a qualitative study in which we conducted structured interviews of physicians and pharmacists from western Nicaragua. The most frequently suggested causes in these discussions and interviews have been urinary tract infection (UTI) and dehydration. Although the appearance of white blood cells in the urinalyses (leukocyturia) is suggestive of a UTI, urine cultures are not a part of the standard medical work-up in Nicaragua. Given the generally infrequent occurrence of UTIs in younger men, caution is advisable before concluding that UTIs are responsible for the frequent occurrence of leukocyturia. As one component of a parallel study in which BU collected blood and urine samples from current ISA workers for analysis of a variety of measures, urine cultures were completed on participants who complained of dysuria or showed evidence of leukocyturia when tested by urine dipstick. If the prevalence of positive urine cultures turns out to be very low, it will be important to pursue other possible explanations for the frequent finding of white blood cells in urinalyses, which could include other types of infections or a tubulointerstitial process. As a result of these preliminary analyses, we have begun abstracting medical records

of additional pilot study subjects in order to have more robust data for analysis. We plan to have a report on these data prepared in 2012.

Regardless of decisions taken by the stakeholders concerning the role of a retrospective cohort study within the Dialogue Table process, BU is interested in finding ways to undertake this and any other activity that will help find the causes of the epidemic of CKD in Western Nicaragua and elsewhere in Central America.

X. Tables

Table 3 Available Demographic Characteristics By Job Category							
Characteristic	Cane Cutters N=58	Cane Gatherers N=29	Pesticide Applicators N=26	Field Machine Operators N=29	Irrigators N=31	Factory Workers N=29	Total N=202
Current Age (yrs)							
N	58	29	26	29	31	29	202
Range	21-71	22-51	19-55	26-64	22-61	28-74	19-74
Median (IQR)	36.9 (30.3, 44.7)	32.9 (27.3, 37.2)	35.0 (29.0, 39.6)	38.2 (33.4, 45.4)	40.0 (31.0, 44.7)	40.2 (35.3, 48.1)	36.9 (30.7, 44.5)
Mean (SD)	38.4 (10.4)	33.2 (7.7)	34.7 (8.0)	39.3 (9.1)	39.3 (9.7)	43.9 (11.7)	38.2 (10.1)
Age at Hire (yrs)							
N	57	29	26	29	31	29	201
Range	14-35	13-40	13-35	12-35	14-57	14-48	12-57
Median (IQR)	19.7 (17.3, 23.5)	19.9 (17.3, 22.3)	21.0 (18.1, 23.4)	20.0 (17.4, 23,2)	19.6 (17.0, 23.9)	21.2 (18.5, 23.3)	20.1 (17.4, 23.3)
Mean (SD)	20.9 (5.0)	21.2 (5.7)	21.7 (4.8)	21.1 (5.3)	21.8 (8.0)	22.4 (7.4)	21.4 (6.0)
Body Mass Index							
N	57	29	26	28	31	29	200
Underweight	4 (7.0%)	1 (3.5%)	1 (3.9%)	1 (3.6%)	1 (3.2%)	1 (3.5%)	9 (4.5%)
Normal weight	45 (79.0%)	19 (65.5%)	18 (69.2%)	9 (32.1%)	25 (80.7%)	12 (41.4%)	128 (64.0%)
Overweight	8 (14.0%)	7 (24.1%)	4 (15.4%)	13 (46.4%)	3 (9.7%)	11 (37.9%)	46 (23.0%)
Obese	0 (0.0%)	2 (6.9%)	3 (11.4%)	5 (17.9%)	2 (6.5%)	5 (17.2%)	17 (8.5%)
Range	16.8, 29.3	17.7, 35.5	18.2, 31.4	18.2, 36.7	18.0, 31.9	18.0, 43.9	16.8,43.9
Median (IQR)	22.3 (20.6, 23.8)	23.4 (20.7, 25.1)	22.9 (20.6, 25.7)	27.1 (22.7, 29.0)	23.2 (20.7, 24.8)	25.7 (21.3, 28.3)	23.2 (21.1, 26.3)
Mean (SD)	23.4 (2.9)	23.7 (3.9)	23.7 (3.7)	26.6 (4.5)	23.0 (3.1)	26.0 (5.8)	24.0 (4.2)

Table 4 Employment Characteristics by Job Category							
Characteristic	Cane Cutters N=91 N (%)	Cane Gatherers N=32 N (%)	Pesticide Applicators N=27 N (%)	Field Machine Operators N=32 N (%)	Irrigators N=32 N (%)	Factory Workers N=32 N (%)	Total N=246 N (%)
Calendar Year Began Employment at ISA							
1960-79	5 (5.5)	0 (0)	1 (3.8)	2 (6.3)	2 (6.3)	6 (18.8)	16 (6.5)
1980-84	9 (10.1)	1 (3.2)	2 (7.7)	4 (12.5)	4 (12.5)	3 (9.4)	23 (9.5)
1985-89	5 (5.6)	2 (6.5)	2 (7.7)	9 (28.1)	6 (18.8)	8 (25.0)	32 (13.2)
1990-94	5 (5.6)	2 (6.5)	2 (7.7)	0 (0)	1 (3.1)	1 (3.1)	11 (4.5)
1995-99	20 (22.5)	13 (41.9)	7 (26.9)	10 (31.3)	10 (31.3)	10 (31.3)	70 (28.9)
2000-04	31 (34.8)	9 (29.0)	7 (26.9)	4 (12.5)	5 (15.6)	4 (12.5)	60 (24.8)
2005-10	14 (15.7)	4 (12.9)	5 (19.2)	3 (9.4)	4 (12.5)	0 (0)	30 (12.4)
Total	89	31	26	32	32	32	242
Calendar Year Began Employment in this Job Category							
1960-79	2 (2.3)	0 (0)	0 (0)	1 (3.2)	1 (3.2)	5 (15.6)	9 (3.7)
1980-84	8 (9.1)	0 (0)	1 (3.8)	1 (3.2)	2 (6.5)	0 (0)	12 (5.1)
1985-89	3 (3.4)	0 (0)	0 (0)	7 (22.6)	4 (12.9)	8 (25.0)	22 (9.4)
1990-94	8 (9.1)	1 (3.7)	4 (15.4)	1 (3.2)	3 (9.7)	2 (6.2)	19 (8.1)
1995-99	22 (25.0)	10 (37.0)	3 (11.5)	6 (19.4)	9 (29.0)	12 (37.5)	62 (26.4)
2000-04	31 (35.2)	11 (40.7)	12 (46.1)	7 (22.6)	6 (13.4)	5 (15.6)	72 (30.6)
2005-10	14 (15.9)	5 (18.5)	6 (23.1)	8 (25.8)	6 (19.4)	0 (0)	39 (16.6)
Total	88	27	26	31	31	32	235
Calendar Year Ended Employment at ISA							
1995-99	3 (3.4)	0 (0)	0 (0)	0 (0)	0 (0)	1 (3.1)	4 (1.6)
2000-04	21 (23.6)	2 (6.5)	3 (11.5)	5 (15.6)	5 (15.6)	5 (15.6)	41 (16.9)
2005-09	34 (38.2)	6 (19.4)	5 (19.2)	8 (25.0)	6 (18.8)	6 (18.7)	65 (26.9)
Still Employed*	31 (34.8)	23 (74.2)	18 (69.2)	19 (59.4)	21 (65.6)	20 (62.5)	132 (54.6)
Total	89	31	26	32	32	32	242

Table 4 Employment Characteristics by Job Category							
Characteristic	Cane Cutters N=91 N (%)	Cane Gatherers N=32 N (%)	Pesticide Applicators N=27 N (%)	Field Machine Operators N=32 N (%)	Irrigators N=32 N (%)	Factory Workers N=32 N (%)	Total N=246 N (%)
Calendar Year Ended Employment in This Job Category							
1995-99	3 (3.4)	0 (0)	0 (0)	0 (0)	1 (3.2)	1 (3.1)	5 (2.1)
2000-04	22 (25.0)	9 (33.3)	4 (15.4)	5 (16.1)	7 (22.6)	6 (18.8)	53 (22.6)
2005-09	35 (39.8)	14 (51.9)	15 (57.7)	10 (32.3)	9 (29.0)	7 (21.9)	90 (38.3)
Still Employed*	28 (31.8)	4 (14.8)	7 (26.9)	16 (51.6)	14 (45.2)	18 (56.3)	87 (37.0)
Total	88	27	26	31	31	32	235
Duration of Employment at ISA[$]							
>0-5 years	67 (75.3)	20 (64.5)	14 (53.9)	15 (48.4)	14 (43.8)	10 (31.3)	140 (58.1)
>5-10 years	12 (13.5)	10 (32.3)	10 (38.5)	5 (16.1)	9 (28.1)	8 (25.0)	54 (22.4)
>10 years	10 (11.2)	1 (3.2)	2 (7.7)	11 (35.5)	9 (28.1)	14 (43.8)	29 (19.5)
Total	89	31	26	31	32	32	241
Duration of Employment at ISA[$]							
N	89	31	26	31	32	32	241
Range	0.1 – 17.3	0.2 – 13.8	0.3 – 26.9	0.4 – 19.0	0.5 – 20.3	0.4 – 31.3	0.1 – 31.3
Median (IQR)	2.1 (0.8, 4.9)	3.2 (1.8, 5.7)	4.5 (2.3, 7.0)	5.2 (2.4, 13.5)	5.9 (2.8, 11.0)	8.0 (4.6, 14.6)	4.1 (1.7, 7.7)
Mean (SD)	3.7 (4.0)	3.8 (2.9)	5.5 (5.4)	7.4 (5.8)	7.2 (5.5)	10.8 (8.0)	5.8 (5.7)
Duration of Employment in this job category[$]							
>0-5 years	72 (81.8)	27 (100.0)	22 (84.6)	18 (62.1)	19 (61.3)	13 (40.6)	171 (73.4)
>5-10 years	16 (18.2)	0 (0)	3 (11.5)	3 (10.3)	9 (29.0)	8 (25.0)	39 (16.7)
>10 years	0 (0)	0 (0)	1 (3.9)	8 (27.6)	3 (9.7)	11 (34.3)	23 (9.9)
Total	88	27	26	29	31	32	233
Duration of Employment in this job category[$]							
N	88	27	26	29	31	32	233
Range	0.1 – 9.6	0.2 – 4.8	0.2 – 13.0	0.4 – 18.8	0.3 – 17.6	0.4 – 25.1	0.1 – 25.1
Median (IQR)	1.8 (0.6, 3.7)	1.4 (0.8, 2.5)	2.6 (0.5, 4.2)	3.3 (1.9, 10.5)	4.2 (1.2, 7.3)	6.6 (4.1, 12.0)	2.5 (0.9, 5.4)
Mean (SD)	2.7 (2.6)	1.8 (1.2)	2.9 (3.0)	6.0 (5.5)	4.8 (4.3)	8.5 (6.4)	4.1 (4.5)

Table 4 Employment Characteristics by Job Category							
Characteristic	Cane Cutters N=91 N (%)	Cane Gatherers N=32 N (%)	Pesticide Applicators N=27 N (%)	Field Machine Operators N=32 N (%)	Irrigators N=32 N (%)	Factory Workers N=32 N (%)	Total N=246 N (%)
Average Percent of Year Employed at ISA[&] N Range Median (IQR) Mean (SD)	 52 16.7 – 72.5 40.6 (33.3, 49.4) 40.4 (13.9)	 31 25.0 – 94.2 47.7 (41.7, 50.0) 47.6 (11.0)	 26 20.8 – 96.7 56.3 (43.5, 70.0) 56.2 (19.1)	 31 25.0 – 93.3 54.5 (46.7, 68.9) 57.7 (16.7)	 32 8.3 – 100.0 55.1 (50.0, 59.8) 56.1 (18.0)	 30 29.2 – 90.9 50.0 (48.5, 57.6) 55.0 (13.5)	 202 8.3 – 100.0 50.0 (41.7, 58.3) 50.9 (16.7)
Average Percent of Year Employed in Selected Job[&] N Range Median (IQR) Mean (SD)	 43 16.7 – 75.0 36.1 (29.2, 41.7) 37.2 (12.4)	 26 16.7 – 51.7 45.5 (37.5, 50.0) 42.4 (8.9)	 24 8.3 – 84.2 38.9 (29.2, 53.7) 42.9 (20.0)	 29 25.0 – 93.3 59.2 (50.0, 80.6) 63.1 (18.6)	 29 33.3 – 100.0 56.7 (50.0, 59.4) 57.5 (14.1)	 30 33.3 – 59.2 52.8 (50.0, 55.6) 52.5 (5.0)	 181 8.3 – 100.0 50.0 (36.1, 56.7) 48.7 (16.6)
Mean Number of Months Employed per Year at ISA[&] N Range Median (IQR) Mean (SD)	 52 2.0 – 8.7 4.9 (4.0, 5.9) 4.9 (1.7)	 31 3.0 -11.3 5.7 (5.0, 6.0) 5.7 (1.3)	 26 2.5 - 11.6 6.8 (5.2, 8.4) 6.8 (2.3)	 31 3.0 – 11.2 6.6 (5.6, 8.3) 6.9 (2.0)	 32 1.0 - 12.0 6.6 (6.0, 7.2) 6.7 (2.2)	 30 3.5 – 10.9 6.0 (5.8, 6.9) 6.6 (1.6)	 202 1.0 – 12.0 6.0 (5.0, 7.0) 6.1 (2.0)
Mean Number of Months Employed per Year in Selected Job[&] N Range Median (IQR) Mean (SD)	 43 2.0 – 9.0 4.3 (3.5, 5.0) 4.5 (1.5)	 26 2.0 – 6.2 5.5 (4.5, 6.0) 5.1 (1.1)	 24 1.0 – 10.1 4.7 (3.5, 6.4) 5.2 (2.4)	 29 3.0 – 11.2 7.1 (6.0, 9.7) 7.6 (2.2)	 29 4.0 – 12.0 6.8 (6.0, 7.1) 6.9 (1.7)	 30 4.0 – 7.1 6.3 (6.0, 6.7) 6.3 (0.6)	 181 1.0 – 12.0 6.0 (4.3, 6.8) 5.8 (2.0)
Ever Held Other Jobs at ISA Yes N (%) Total	 43 (48.3) 89	 28 (90.3) 31	 22 (84.6) 26	 21 (65.6) 32	 22 (68.7) 32	 26 (81.2) 32	 162 (66.9) 242

Table 4 Employment Characteristics by Job Category							
Characteristic	Cane Cutters N=91 N (%)	Cane Gatherers N=32 N (%)	Pesticide Applicators N=27 N (%)	Field Machine Operators N=32 N (%)	Irrigators N=32 N (%)	Factory Workers N=32 N (%)	Total N=246 N (%)
Other Jobs at ISA							
Cane Cutter	--	8 (25.0)	9 (34.6)	6 (18.8)	8 (25.0)	0 (0)	31 (12.8)
Cane Gatherer	5 (5.6)	--	3 (11.5)	6 (18.8)	3 (9.4)	2 (6.2)	19 (7.8)
Pesticide Applicator	27 (30.3)	10 (32.3)	--	5 (15.6)	9 (28.1)	2 (6.2)	53 (21.9)
Field Machine Op.	3 (3.4)	22 (71.0)	6 (23.1)	--	9 (28.1)	5 (15.6)	45 (18.6)
Irrigator	13 (14.6)	6 (19.4)	9 (34.6)	1 (3.1)	--	4 (12.5)	33 (13.6)
Factory Worker	3 (3.4)	0 (0)	1 (3.8)	2 (6.2)	2 (6.2)	--	8 (3.3)
Other[#]	27 (30.3)	14 (45.2)	14 (53.8)	11 (34.4)	17 (53.1)	18 (56.2)	101 (41.7)
Unknown[%]	23 (25.8)	10 (32.3)	13 (50.0)	14 (43.8)	12 (37.5)	15 (46.9)	87 (35.9)
Total	89	31	26	32	32	32	242
Duration of Employment in Other Jobs[$]							
>0-5 years	38 (88.4)	25 (89.3)	20 (90.9)	19 (90.5)	16 (72.7)	21 (80.8)	139 (85.8)
>5-10 years	4 (9.3)	2 (7.1)	0 (0)	2 (9.5)	5 (22.7)	4 (15.4)	17 (10.5)
>10 years	1 (2.3)	1 (3.6)	2 (9.1)	0 (0)	1 (4.6)	1 (3.8)	6 (3.7)
Total	43	8	22	21	22	26	162
Duration of Employment in Other Jobs[$]							
N	43	28	22	21	22	26	162
Range	0.1 – 10.1	0.4 – 12.4	0.4 – 13.9	0.1 – 9.6	0.1 – 12.9	0.0 – 19.1	0.0 – 19.1
Median (IQR)	1.4 (0.3, 3.5)	1.5 (0.8, 3.4)	2.1 (1.0, 3.7)	1.8 (1.1, 3.2)	2.7 (1.0, 5.5)	1.1 (0.6, 2.3)	1.6 (0.7, 3.7)
Mean (SD)	2.2 (2.4)	2.5 (2.5)	3.0 (3.2)	2.6 (2.4)	3.8 (3.3)	2.7 (4.2)	2.7 (3.0)

Table 4 Employment Characteristics by Job Category							
Characteristic	Cane Cutters N=91 N (%)	Cane Gatherers N=32 N (%)	Pesticide Applicators N=27 N (%)	Field Machine Operators N=32 N (%)	Irrigators N=32 N (%)	Factory Workers N=32 N (%)	Total N=246 N (%)
Longest Job Held at ISA							
Cane Cutter	79 (88.8)	1 (3.2)	0 (0)	1 (3.2)	2 (6.2)	0 (0)	83 (34.4)
Cane Gatherer	2 (2.2)	16 (51.6)	1 (3.8)	1 (3.2)	0 (0)	0 (0)	20 (8.3)
Pesticide Applicator	2 (2.2)	2 (6.5)	16 (61.5)	0 (0)	1 (3.1)	0 (0)	21 (8.7)
Field Machine Op.	0 (0)	10 (32.3)	1 (3.8)	24 (77.4)	2 (6.2)	0 (0)	37 (15.4)
Irrigator	2 (2.2)	2 (6.5)	5 (19.2)	0 (0)	24 (75.0)	0 (0)	33 (13.7)
Factory Worker	0 (0)	0 (0)	1 (3.8)	1 (3.2)	0 (0)	29 (90.6)	31 (12.9)
Other[#]	3 (3.4)	0 (0)	1 (3.8)	3 (9.7)	1 (3.1)	2 (6.2)	10 (4.2)
Unknown[%]	1 (1.1)	0 (0)	1 (3.8)	1 (3.2)	2 (6.2)	1 (3.1)	6 (2.5)
Total	89	31	26	31	32	32	241
Mean # days worked in a 14 day period[&]							
N	52	31	26	31	32	30	202
Range	3.7 – 7.0	5.7 – 7.0	5.4 – 6.9	6.1 – 7.0	2.0 – 7.0	6.8 – 7.0	2.0 – 7.0
Median (IQR)	6.4 (6.1, 6.7)	6.8 (6.6, 6.9)	6.6 (6.3, 6.7)	6.9 (6.8, 7.0)	6.9 (6.8, 7.0)	7.0 (7.0, 7.0)	6.8 (6.6, 6.9)
Mean (SD)	6.3 (0.7)	6.7 (0.3)	6.5 (0.4)	6.9 (0.2)	6.7 (0.9)	6.9 (0.1)	6.6 (0.6)
Mean # number of hours worked in a 14 day period[&]							
N	NA[@]	31	26	31	32	30	202
Range		13.9 – 78.9	7.4 – 74.7	46.0 – 76.3	11.5 – 78.9	58.0 – 75.7	3.8 – 80.8
Median (IQR)		72.5 (68.1, 75.9)	53.0 (36.4, 57.0)	70.1 (62.9, 74.7)	69.0 (60.9, 76.1)	71.8 (67.7, 73.0)	65.1 (37.2, 72.2)
Mean (SD)		67.9 (14.5)	48.8 (15.6)	67.4 (7.9)	63.9 (16.4)	70.4 (4.2)	54.0 (24.1)

* Employed in 2010
[$] Duration is calculated as of 2010
[&] Based only on electronic records
[#]Includes other types of machine operators, guards, shopkeepers, administrative positions
[%] These job types could not be determined and may include selected jobs
[@] Not available for contracted cane cutters.

Table 5 Medical Record Linkage Rates at ISA Hospital by Job Category		
Job Category	# Medical Records Found/ # In Job Category	Linkage Rate (%)
Cane Cutters	64/90	71%
Cane Cutters during: 1997-1999	20/21	95%
2000-2002	19/20	95%
2003-2006	13/28	46%
2007-2010	12/21	57%
Irrigators	30/32	94%
Pesticide Applicators	25/26	96%
Field Machine Operators	32/32	100%
Factory Workers	32/32	100%
Cane Gatherers (Puchos)	31/31	100%
Total	214/243	88%

Table 6: Annual Number of Medical Encounters by Selected Job			
		Number of encounters per worker per year	
Job Title	# of Workers	Range	Median
Cane Cutter	15	1.0 - 5.0	1.8
Irrigator	9	1.7 - 4.0	2.0
Pesticide Applicator	9	1.3 - 4.6	2.0
Field Machine Operator	9	1.2 - 5.2	1.7
Factory Worker	10	1.2 - 4.8	1.5
Cane Gatherer (Pucho)	9	1.0 - 2.3	1.4
Total	61	1.0 - 5.2	1.7

Table 7: Number of Medical Encounters Over Time				
Year of Encounter	Number of Workers	Number of Encounters	Range	Median
1965-1969	1	2	--	2.0
1970-1974	2	6	1-5	3.0
1975-1979	2	15	1-14	7.5
1980-1984	7	16	1-4	2.0
1985-1989	14	58	1-25	2.5
1990-1994	22	96	1-36	2.0
1995-1999	32	200	1-32	3.0
2000-2004	51	360	1-28	6.0
2005-2009	41	308	1-28	6.0
2010	27	55	1-9	2.0
Unknown	2	2	1-1	1.0

Table 8: Frequency of Medical Diagnoses		
Diagnostic Category	# of Encounters	% of Encounters
Arthropod bites & stings	1	0.1
Dermatologic diseases	20	1.8
Diseases of the respiratory system	42	3.8
Disorders of eyes, ears, nose, and throat	43	3.9
Disorders of the cardiovascular system	17	1.5
Disorders of the gastrointestinal system	44	4.0
Disorders of the immune system	3	0.3
Disorders of the joints & adjacent tissues	102	9.1
Disorders of the kidney & urinary tract	174	15.6
Endocrinology & metabolism	14	1.3
Hematologic alterations	6	0.5
Infectious diseases	233	20.8
Neoplastic disorders	1	0.1
Neurologic disorders	4	0.4
Poisoning & drug overdose	9	0.8
Psychiatric disorders	3	0.3
Trauma & burns	47	4.2
Uncategorized	47	4.2
General Medical Examination	244	21.8
No Diagnosis Listed	190	17.0
Total encounters	1118	100.0

Table 9: Frequency of Urinary Tract Infection Diagnosis by Job Category		
	Frequency of Diagnosis	
	# of Workers	% of Workers
Cane Cutter	12	80.0
Irrigator	8	88.9
Pesticide Applicator	6	66.7
Filed Machine Operator	6	66.7
Factory worker	6	60.0
Cane Gatherer (Pucho)	4	44.4
Total	42	100.0

Table 10: Urinalysis Results Among Workers	
Urinalysis Component	# of Workers (%)
Presence of:	
Protein (mg/dL) Trace 30 100 – 500 > 500	 17 (27.9) 15 (24.6) 5 (8.2) 0 (0)
Nitrites	4 (6.6)
White blood cells (>5/hpf)	28 (45.9)
Red blood cells (>3/hpf)	20 (32.8)
Casts Leukocyte Granular Red blood cell	 13 (21.3) 10 (16.4) 0 (0.0)

Table 11: Comparison of CKD Cases by Definition of Diagnosis			
CKD Diagnosis Noted in Medical Record	CKD diagnosis based on two serum creatinine levels of >1.4 mg/dL at least three months apart		
	No	Yes	Total
No	53 (86.9%)	1 (1.6%)	54 (88.5%)
Yes	1 (1.6%)	6 (9.8%)	7 (11.5%)
Total	54 (88.5%)	7 (11.5%)	61 (100.0%)

Table 12: Frequency of CKD* According to Longest Job		
Job Title with Longest Duration	# of Workers with CKD/ # of Workers in Job Category	% of Workers
Cane Cutter	5/17	29.4
Filed Machine Operator	1/10	10.0
Factory Worker	1/11	9.1
Irrigator	0/7	0.0
Pesticide Applicator	0/7	0.0
Cane Gatherer (Pucho)	0/6	0.0
Other	0/2	0.0
Unknown	0/1	0.0
Total	7/61	11.5

*Creatinine-based CKD definition

Table 13: Frequency of CKD* According to Selected Job Category		
	# of Workers with CKD/ # of Workers in Job Category	% of Workers
Cane Cutter	3/15	20.0
Irrigator	1/9	11.1
Pesticide Applicator	2/9	22.2
Filed Machine Operator	1/9	11.1
Factory worker	0/10	0.0
Cane Gatherer (Pucho)	0/9	0.0
Total	7/61	11.5

* Creatinine-based definition

Table 14: Mean Blood Pressure Levels Among All Workers and Workers with CKD		
	Blood pressure (mm Hg)	
	Systolic	Diastolic
All Workers (SD)	112.8 (10.0)	72.5 (7.3)
Workers with CKD (SD)	114.4* (10.2)	69.7* (10.2)

* Before CKD diagnosis

Table 15 Most Commonly Prescribed Medications by Medication Class		
Medication Class	# of Encounters	% of Encounters
NSAID	280	25.1
Analgesic/antipyretic	17	1.5
Analgesic/antipyretic + decongestant + antihistamine	115	10.3
Antacid	30	2.7
Anti-asthmatic	4	0.4
Antibiotic	374	33.5
Antiemetic	6	0.5
Antifungal	9	0.8
Antihistamine	48	4.3
Antihypertensive	11	1.0
Antimalarial	16	1.4
Antiparasitic	12	1.1
Antispasmodic + analgesic	12	1.1
Antitussive	10	0.9
Antitussive + antihistamine + expectorant + decongestant	23	2.1
Anxiolytic	10	0.9
Bronchodilator	6	0.5
Mucolytic	32	2.9
Other	61	5.5
Steroid	12	1.1
Topical medication	10	0.9
Urinary analgesic + urinary antiseptic	22	2.0
Vitamin	83	7.4
Drug used in gout	12	1.1

Table 15 Most Commonly Prescribed Medications by Medication Class		
Medication Class	# of Encounters	% of Encounters
Drugs used in diabetes	4	0.4
Total Encounters	1118	100.0

Table 16 Demographic and Lifestyle Characteristics of Subjects With In-Person Interviews			
	Current Workers (N=5)	Former Workers (N=5)	Total (N=10)
Age Range (years)	28-48	32-70	28-70
Currently Married	0%	60%	30%
Drinks Tap or Bottled Water	100%	40%	70%
Drinks Well Water	0%	60%	30%
Drinks Alcoholic Beverages	100%	80%	90%
Ever Drank Lija	60%	20%	40%
Ever Smoked Cigarettes	40%	60%	50%
Ever Raised Crops for Food	40%	40%	40%
Ever Used Pesticides on Crops	40%	20%	30%
Ever Raised Animals* for Food	20%	100%	60%

* Mainly chickens

Table 17 Work Characteristics of Subjects With In-Person Interviews			
Characteristic	Current Workers (N=5)	Former Workers (N=5)	Total (N=10)
Ever Employed as a Cane Cutter at ISA	40%	80%	60%
Ever Employed as a Herbicide or Pesticide Applicator at ISA	40%	60%	50%
Ever Employed as a Machine Operator at ISA	*60%*	*0%*	30%
Ever Employed as an Irrigator at ISA	20%	60%	40%
Ever Employed as Cane Gatherer at ISA	20%	0%	10%
Ever Employed as Factory Worker at ISA	0%	20%	10%
Ever Employed in Another Job at ISA*	40%	80%	60%
Number of Jobs Held at ISA* 1 2-3 4-5 >5	 60% 0% 20% 20%	 0% 60% 20% 20%	 30% 30% 20% 20%
Ever Employed as a Cane Cutter outside of ISA	0%	20%	10%
Number of Jobs Held Outside of ISA** 0 1 2-3	 40% 40% 20%	 60% 40% 0%	 50% 40% 10%
Exposed to Pesticides or Herbicides at ISA	80%	60%	70%
Exposed to Pesticide or Herbicide Outside of Work	20%	20%	20%
Exposed to Dust at ISA	80%	100%	90%
Exposed to Dust Outside of Work	40%	0%	20%
Exposed to Solvents at ISA	20%	0%	10%
Exposed to Solvents Outside of Work	0%	0%	0%
Exposed to Lead, Cadmium and Gasoline at ISA or Outside of Work	0%	0%	0%

*Includes grounds maintenance, guard during field burning, weeder, seeder, driver, and security guard

** Includes driver, pest control, brick layer, farmer and security guard

Table 18 Medical Characteristics of subjects with In-Person Interviews			
History of :	Current Workers (N=5)	Former Workers (N=5)	Total (N=10)
CRI	0%	80%	40%
Kidney Infection	40%	100%	70%
Chistata	40%	100%	70%
Dehydration	40%	40%	40%
Hypertension	40%	40%	40%
Diabetes	20%	0%	10%
Pesticide Poisoning	20%	0%	10%
Heat Stroke	40%	20%	30%
Blood in Urine	20%	40%	30%
Protein in Urine	0%	0%	0%
Elevated Creatinine	20%	100%	60%
Abnormal Kidney Ultrasound	0%	60%	30%
Gentamicin Use	40%	20%	30%
NSAID Use	80%	100%	90%
Folk Medicine Use	20%	20%	20%
Family Member with CKD	60%	60%	60%

Table 19 Medical Care Facilities Used by Subjects with In-Person Interviews			
	Current Workers (N=5)	Former Workers (N=5)	Total (N=10)
ISA Hospital	100%	100%	100%
Hospital Espana	20%	40%	30%
Hospital Escuela Oscar Danilo Rosales (HEODRA)	40%	40%	40%
Centro de Salud in Chichigalpa	40%	80%	60%
Other*	80%	0%	40%

* Includes Hospital Mauricio Abdalah in Chinandega, Amorsa in Chichigalpa, a private doctor in Leon, and “la consulto” in Chinandega

Appendix: Detailed Methods for Selection of Pilot Study Subjects

1. Cane Cutters
   Seven came cutters per year (n=90, due to one person being selected twice) were selected. When possible, cane cutters were selected from a February payroll, in the middle of the zafra, to maximize the likelihood that the population of workers represented cane cutters who worked during most of the zafra (excluding those that stopped working very early and perhaps those that began work very late).

   During zafras 1997, 1998 and 1999: only paper ISA payroll records were available. Cane cutters are listed in Planillas 10 and 11, which were alternated as the source for selection in each successive year as was the first or second payment period in the month. Within the planillas, records are ordered by ficha. Seven workers were systematically selected (i.e., every n^{th} record) from lower to higher ficha, excluding the first and last two full pages of workers so as to avoid the lowest and highest fichas.

   During zafras 2000, 2001 and 2002: during these years, records were available only in electronic files and only for cane cutters who were ISA employees. All cane cutters in planillas 10 and 11 were provided to us in a file. The cane cutters were sorted by ficha and then seven were systematically selected from lower to higher ficha, avoiding fichas in the very lowest and highest ranges. It was not necessary to alternate planillas for selection, because all records were in the same file.

   Contracted workers do appear in the "orden de pegue" file during this time. ISA fichas were used to identify workers in 2000 and 2001, and carnets were used in 2002. It might be possible to use this file for purposes of selection, but it would require obtaining personal information by linking to either the ISA employment record or an Enganche Database (which would require that a contracted worker work after 2003).

   During zafras 2003-2006: all cane cutters were contracted during this period and the original identifying information was in the Enganche Screening Database maintained by the Office of Responsibilidad Social Salud en el Campo. Even after excluding those who didn't pass the screening, the database contains many more individuals than were hired. We randomly selected records and then looked for that carnet (carnet) in the "orden de pegue" for that year. If the carnet appeared, we considered that confirmation that the person had been hired. If a carnet did not appear, we went to the next record and used the same procedure. Worker population from which subjects were selected represented all workers that were hired into the zafrat (i.e. could not replicate process of selecting in February). Over the four years, we had to identify 48 people in order to confirm 28. In zafras 2003 through 2005, supposedly only cane cutters were screened. However, 77 applicants in 2003, 231 in 2004, and 75 in 2005 had the entry "Fabrica" in the Contrato column. There were also 7 applicants in 2003 with the entry "Riego" in the Contrato column. None of these workers were selected in any year. In zafra 2006, other contracted field workers were also included in the screening. They were not distinguished from the cane cutters in the file. However, only cane cutters would appear in the "orden de pegue".

   The contractor listed in the pre-screening file was not necessarily the same as in the "orden de pegue" and contractors for a particular worker could also change within the "orden de pegue" in the same year. ISA staff reported that this type of switch does occur, perhaps because a particular contractor and worker don't get along or other issues.

In the 2004 file only, one column with no header had the following entries for different records: impreso (literally, “printed”; 666 people), ponchado (“struck”; 951), base de datos (107), Nuevo (1566), fisico (1), and hacer (4). The 951 people labeled “ponchado” all had the status of “Si”.

During zafras 2007 through 2009: It was not necessary to use the “orden de pegue” file, because contracted cane cutters were captured in a specific payroll maintained by an ISA staff member. We were able to select cane cutters by ordering the payroll by carnet and selecting systematically as for 2000 through 2002. Payroll records consisted of pdf files of payroll organized by Contractor (contratista), and ordered from low to high contractor carnet within each file. Files are produced every week pay period throughout the zafra. A different week in February was chosen for each zafra season (’07-’08: Week 14, ’08-’09: Week 13, ’09-’10: Week 15). All contractor payroll files were combined for that week and then 7 workers were chosen at random from the combined file. Many files had INSS number included, but for those that did not, INSS number and any other available information were retrieved from the Enganche Screening Database, if available there.

2. Irrigators
   2-3 irrigators were selected from each year for a total of 32. Irrigators were selected from Planilla 5 during the month of January.

   For zafra 1997 and zafra19 98: Irrigators were selected from Codigo Processo with the description: Nucleo XX, Reigo y Drenaje, Zona No.1 and Zona No.2. Random pages were selected from the all of the pages with these Codigo Processo. On a selected page, a random worker was chosen and selected if he had a job title of Operador General de Campo.

   For zafra 1999-00: Irrigators were selected from Codigo Processo 2151- Reigo y Drenaje using the same method as described above. Job Title: Operador General de Campo. This method was also used for the remaining job categories.

   For zafra 00-01 and zafra 01-02: We received the entire planilla 5. The CODCAT (641) and job title description (Operario de Campo) did not specifically identify irrigators. We selected those with CODPRO=2151 (Riego y Drenaje). We ordered each year by ficha and systematically selected either two (in odd-year zafras) or three (even-year zafras) per year. The same approach was used for each of the remaining job categories.

   For zafra 2002 through 2009:. CODCAT=188 (specific job code) was sufficient to identify irrigators. We ordered each year by ficha and systematically selected either two (in odd-year zafras) or three (even-year zafras) per year. The same approach was used for each of the remaining job categories.

3. Applicators
   2-3 irrigators were selected from each year (excluding zafras 1997 and 1998, see below) for a total of 26, due to one person being selected twice. Irrigators were selected from Planilla 5 during the month of June.

   For zafra 1997 and zafra 1998: We were unable to select applicators. There did not appear to be the quantity of pesticide applicators in these years that would have been necessary for selection, or would have been expected for work (only about 15 were listed).

We chose not to pursue other means of identification unless it appeared that we would be proposing these years to select workers for the study.

For zafra 1999: Applicators were selected from Codigo Processo 2161: Control de Maleza and 2171: Coltrol de Plagas, and from Job Title: Operador General de Campo

For zafra 2000 and zafra 2001: We received the entire planilla 5. The CODCAT (641) and job title description (Operario) did not specifically identify applicators. We selected those with CODPRO=2160 (Control de Maleza) or 2171 (Control de Plagas). The latter represents pesticides as opposed to herbicides, and comprised about 1/3 of the total. Selection was by ficha in the usual manner.

For zafra 2002 through 2009: There were supposed to be two "codigo funcional" requested: 174 (herbicides) and 204 (pesticides), but through a miscommunication only 174 was requested. The code 174 was specific from 2002 through 2009, so this was all we received during that time period.

4. Field Machine Operators (including mechanics, drivers, welders, carpenters, etc). Machine operators in the factory were not included in this category.
   2-3 field machine operators were selected from each year for a total of 32. Operators were selected from Planilla 4 during the month of February. There was no simple way to select the jobs of interest from the entire list because the CODCAT and CODPRO varied across years. We reviewed the planillas for 2000-2002, when electronic files first became available, and selected jobs based on the text description of the job title. The majority of workers in the planilla fell into the job categories of interest. For example, in 00-01 there were 1582 workers in planilla 4 of whom 1099 were considered to be in a job of interest.

   For zafra 1997 through zafra 1999: Field machine operators were selected from Planilla 4 using the list of job titles selected from years 2000-2002.

   For zafra 2000 and zafra 2002: Selection was by ficha in the usual manner.

   For zafra 2003 through 2009: Machine operators in the factory were also not included. For zafras 2003 through 2009, the requested codes were sufficient to identify the jobs of interest.

5. Factory Workers
   2-3 factory were selected from each year for a total of 32. Factory workers were selected from Planillas 12 and 13 during the month of February. Because we were interested in factory workers who carried out production tasks, as opposed to technicians and managers, we reviewed the planillas for 2000-2002 as for field machine operators, and selected jobs based on the text description of the job title. The majority of workers fell into the job categories of interest. For example, in 2000 there were 348 workers in planillas 12 and 13 of whom 226 were considered to be in a job of interest.

   For zafra 1997 through zafra 1999: Factory workers were selected from all of Planilla 12 or 13 (alternating across years). In 99-00, only the following Codigo Processos were included: 4121: Extraccion de Jugo, 4141, Produccion Azucar, 4321: Generación Vapor.

   For zafra 2000 and zafra 2001: Selection was by ficha in the usual manner.

For zafra 2003 through 2009: The requested codes were sufficient to identify the jobs of interest.

**One important point to note is that factory workers who work during the non-zafra period are paid under planilla 4, even if they worked during zafra and were paid in planillas 12 and 13. This is simply because there are so few remaining that it is administratively simpler just to integrate them into planilla 4. This did not affect the selection strategy because we were selecting from February payrolls.*

6. Puchos
2-3 puchos were selected from each year for a total of 31, due to one person being selected twice. Puchos were selected from Planilla 4 during the month of February.

For zafra 1997 and zafra 1998: Puchos were selected from Codigo Processo IKW with the description: Area de Puchos. Random pages were selected from the all of the pages with these Codigo Processo. On a selected page, a random worker was chosen and selected if he had a job title of Cade/Gavio 4 Zaf/10

For zafra 1999: Puchos were selected from Codigo Processo 2231: Acarreo using the same method as described above. Job Title: Cade/Gavio.

For zafra 2000 through zafra 2002: For zafra 2000, we selected puchos from the CODPRO 2231 and CODCAT 813:"cadenero." An ISA staff member explained that a cadenero refers to a person who receives the cane from the field at the factory and transports it into the factory by use of an overhead chain. In 2000, there were 247 cadeneros, which was more than in any subsequent year, when the job titles became more differentiated. It is likely that the description cadenero included titles that appeared in later years including gaviota and limipador de trailers, as well as pucho. Because these other job categories may have had similar exposures to puchos, for purposes of the pilot study we included all CADENERO in the pucho job category. In 2001 and 2002, a CODCAT for puchos appeared, and the number of cadeneros decreased (21 in 2001 and 22 in 2002). Jobs were selected from both titles in these years.

In zafra 2001 and 2002, the CODCAT numbers changed to 827 and 241, respectively, but the accompanying description (PUCHO DE ACARREO) was explicit and in both years the CODPRO 2231 was appropriate for puchos; therefore, we treated these categories as the puchos for those years. In zafra 2001, the number of records identified as puchos was substantially lower than in any other year, raising the possibility that some puchos may have been missed because they had different descriptions. Longitudinal information from the payroll record may be helpful in clarifying this issue. Other potential explanations are use of contractors for that year or simply a smaller number of hires. The number of workers identified as puchos in zafra 2002 returned to the usual range.

For zafra 2003 through 2009:. In this period, puchos are listed as OPERARIO GENERAL DE COSCHA with CODCAT=235. The CODPRO was not provided for these years as part of the original request, and we did not think it was necessary to request that it be added.