THE MEANINGFUL USE OF INCORPORATING OCCUPATION AND INDUSTRY IN AN ELECTRONIC HEALTH RECORD (EHR) DATABASE FOR EPIDEMIOLOGIC RESEARCH

by

CHRISTINA MARIA SOCIAS, AB, MPH

APPROVED:

SHARON COOPER, PHD

GEORGE DELCLOS, MD, PHD

JULIE GRAVES, MD, PHD

YUANYUAN LIANG, PHD

DEAN, THE UNIVERSITY OF TEXAS
SCHOOL OF PUBLIC HEALTH

DEDICATION

To my father, Jesus Socias, the hardest worker I'll ever know.

THE MEANINGFUL USE OF INCORPORATING OCCUPATION AND INDUSTRY IN AN ELECTRONIC HEALTH RECORD (EHR) DATABASE FOR EPIDEMIOLOGIC RESEARCH

by

CHRISTINA MARIA SOCIAS, AB, MPH

Presented to the Faculty of The University of Texas

School of Public Health

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PUBLIC HEALTH

THE UNIVERSITY OF TEXAS
SCHOOL OF PUBLIC HEALTH
Houston, Texas
August 2013

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

PREFACE

Everyday millions of people leave their home and family to earn a living. While most return home just as they left, a number become sick, injured or even die due to a hazardous workplace. Occupational Epidemiology is the study of working populations to reduce and prevent illness, injury or death. As an upcoming Occupational Epidemiologist, I aspire to see the day where *all* workplaces become safe and healthy.

ACKNOWLEDGEMENTS

Thank you to my advisor, mentor, and close friend, Sharon Cooper. You have been an incredible blessing in my life academically and personally. Thank you, George Delclos, for your expertise and incredible patience. I've learned so much from both of you over the years and aspire to be half the dedicated mentors you have been to me! Yuanyuan Liang and Julie Graves, although this was our main project together, I couldn't have done it without your individual knowledge of statistics and EHRs, thank you both for your enthusiasm and support. Together, you were the dream team!

Thank you to the Community Based Research Network (CBRN) - Bobbi Ryder, Nick Heyer, Ed Hendrikson, Sharon Cooper, Eva Shipp, Deborah Del Junco, Melissa Valerio, Alex Zottarelli, and Nicole Henry- for caring about the Farmworkers and their families as though they were your own family. Thank you specifically to Ed Hendrikson and Maria DeJesus at Salud Family Health Center for sharing your data and consulting on this pioneering effort.

Thank you to the data management experts- Eva Shipp, Charlie Cooper and David Tucker. Without your great skills, I would still be running queries in Microsoft Access. Thank you to Nancy Weiss and Pam Schumacher at the NIOSH Division of Surveillance, Hazard Evaluations, and Field Studies for teaching me about the complexities of coding industry and occupations. Thank you to Elaine Symanski and Sarah Felknor for supporting the work of doctoral students in occupational and environmental health. Your hard work pays off directly

to us. Thank you to the very patient administrative staff at UTSPH- specifically Melissa Garcia and Rebecca Novak.

Funding and research support was provided by grant number T42CCT610417 from the National Institute for Occupational Safety and Health (NIOSH)/Centers for Disease Control and Prevention (CDC) to the Southwest Center for Occupational and Environmental Health (SWCOEH), a NIOSH Education and Research Center (ERC) and grant number RC4ES019405 from the National Institute of Environmental Health Sciences (NIEHS).

Thank you to God, my family, and my friends for spiritual and emotional support, even when I thought I didn't need it. Thank you to Gustavo Morales for supporting me through the stress, the late nights and the crazy times. Thank you to my family- my blessed parents, my incredibly intelligent siblings (you're next!), my grandparents, and my uncle. Thank you to my friends- Chaz, Ruth, Robyn, Laura, Amery, and to everyone who has cheered me on!

THE MEANINGFUL USE OF INCORPORATING OCCUPATION AND INDUSTRY IN AN ELECTRONIC HEALTH RECORD (EHR) DATABASE FOR EPIDEMIOLOGIC RESEARCH

Christina Maria Socias, AB, MPH, DrPH
The University of Texas
School of Public Health, 2013

Dissertation Chair: Sharon Cooper, PhD

Introduction: Few extensive, national clinical databases exist on the health of migrant and seasonal farmworkers (MSFWs). Electronic health records (EHRs) are increasingly utilized by Federally Qualified Health Centers (FQHCs) and have the potential to improve clinical care, as well as supplement current surveillance and epidemiologic studies of underserved populations, such as MSFWs. The goal of this study was to examine the feasibility of using an EHR database in a descriptive epidemiologic study of a vulnerable working population, MSFWs. The aims of this study were to describe the patients at a Colorado FQHC; to assess the feasibility of coding patient occupation using the National Institute of Occupational Safety and Health (NIOSH) Industry and Occupation Computerized Coding System (NIOCCS) and to describe cholinesterase values from a subset of patients, as a marker of pesticide exposure.

Methods: Patients were described by MSFW status on demographics, social history, medical indicators, and diagnoses from a de-identified EHR database from a large, multi-site Colorado Migrant Health Center (MHC). Logistic regression models were constructed for hypertension diagnosis and separately for elevated blood pressure at the last clinic visit. Laboratory cholinesterase values were compared from baseline to exposure period.

Results: 41,817 patients from 2012 were included in the study: 553(1.3%) MSFWs, 20,665(49.4%) non-MSFWs and 20,599(49.3%) had no information in the MSFW field. MSFWs were more often male, married, employed, Hispanic, and Spanish speaking compared to Non-MSFWs. The most frequent diagnoses for all patients were hypertension, overweight/obesity, lipid disorder, type 2 diabetes, back disorder. Risk factors for hypertension included mainly age, obesity, and sex, but did not include MSFW status. Laboratory cholinesterase values showed a decrease from baseline to after suspected exposure in the 12 patients.

Discussion: Although there were significant missing values, this study was able to analyze medical data in a timely manner and show that meaningful use requirements can improve the usability of EHR data for epidemiologic research of MSFWs and other patients at FQHCs. The results of the study were consistent with current literature available for MSFWs. This innovative data source may be the next major development for occupational injury and illness surveillance and research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

BACKGROUND: INTRODUCTION

Occupation and Health Outcomes History in America

Workplace illness and injury investigations are long established endeavors. However, workers still suffer illness and injury, and may benefit from innovative methods to explore the workplace exposure-outcome relationship. An Italian physician, Bernardino Ramazzini (1663-1714), introduced the idea of a connection between occupation and health over three centuries ago (1,2). Often cited as the "father of occupational medicine," he described many occupational related diseases and causes in 1700 (2). More recently, several movements and disasters have encouraged safer and healthier workplaces around the world and specifically in the United States (U.S.), including Alice Hamilton's prevention efforts during World War I, the Triangle Shirtwaist Fire of 1911, the Farmington Coal Disaster of 1968, and other preventable disasters. Similar to other epidemiology sub-disciplines, occupational epidemiology arose from several case-series investigations and from keen medical providers curious about rare diseases in small, specific populations (1,2).

In an effort to reduce the occurrence of such injury and illness, the U.S. Congress passed the Occupational Safety and Health (OSH) Act of 1970 to establish the Occupational Safety and Health Administration (OSHA) as part of the Department of Labor. OSHA became the federal agency charged with assuring "safe and healthful working conditions for working men and women by setting and enforcing standards and by providing training, outreach, education and assistance" (3). Further, the OSH Act established the National Institute for Occupational Safety and Health (NIOSH) to be part of the Centers for Disease

Control and Prevention (CDC) in the US Department of Health and Human Services (HHS). The main goal of NIOSH is "to help assure safe and healthful working conditions for working men and women by providing research, information, education, and training in the field of occupational safety and health" (4). Effectively, OSHA functions as the enforcement agency for workplace health and safety, while NIOSH furthers research and education focused on raising expertise in the field (1).

Occupational epidemiology is the study of illness and injury in the workplace including "the systematic collection, analysis, and dissemination of health-related information for purposes of prevention or control of disease or injury, including occupational surveillance beyond mortality and morbidity, to information about injuries, hazards, and exposures"(1,2).More recent issues such as "popcorn lung," pesticide- related illness, latex powder associated allergies, and other occupationally acquired illnesses have been exposed and reduced as a result of federal investigation (1). These investigations were carried out using many of the principles of epidemiology, and today the field of occupational epidemiology serves many purposes at NIOSH, OSHA and other federal agencies. In spite of the progress made in occupational health, the burden of disease and injury remains substantial. A review of occupationally related illness and injury literature using primary and secondary data estimates that over $128 billion is spent on related costs annually in America (5,6). Most people work to subsist and the conditions of their workplace should be safe, healthy, and support both the economy of the individual, businesses and the community.

According to the OSH Act, health and safety are often related to occupation, impacting "interstate commerce in terms of lost production, wage loss, medical expenses, and disability compensation payments" (7). Historically, investigations of occupational illness and injury have been conducted in the workplace as an issue arises that has often involved arduous medical chart reviews and lengthy data entry. With the upcoming modernization of health records to usable, interoperable electronic health records (EHRs), medical professionals may, again, be at a unique position to improve worker health and understanding of a patient's occupational risk through EHR data. EHR databases and Health Information Exchange (HIE) databases have been targeted as an efficient and effective resource to utilize for public health alerts, surveillance, and research, mainly of infectious diseases (8,9). A HIE is defined as the transmission and sharing of patient-level health data between providers (10,11). An EHR database is a collection of individual patient records from one or more providers and a HIE database is a collection of EHR data shared between providers. Such databases may also assist clinicians and public health professionals in understanding the risks faced by certain underserved working populations where occupation and industry information is available and traditional data sources are difficult to obtain, such as is the case with migrant farmworkers (12). Further, incorporating such occupation and industry variables may be vital to realizing the full potential of federally-defined meaningful use standards (13), specifically during the upcoming stage 3 which focuses on population health.

Farmworkers in America

Farmworkers represent a substantial number of American workers with estimates ranging from approximately 750,000 (14) to over three million (15). Most farmworkers are seasonal and/or migrant workers, and are diverse in nationality, socioeconomic status, and country of birth. According to a recent survey, approximately 81% of farmworkers are Hispanic or Latino with about 75% born in Mexico, and an estimated 53% are unauthorized to work in the U.S. This survey found that the average income of a farmworker family is between $15,000 and $17,499 and only 23% have health insurance (16). Due to their unique laborious working situations and variety of backgrounds, farmworkers have an increased risk of injury and adverse health outcomes, and experience many barriers to healthcare access such as literacy, language, financial or even transportation (15,16).

An estimated 862,808 migrant and seasonal farmworkers and their family members received care at a migrant health center in 2011 (17). Clinical data acquired at MHCs may be an important source for surveillance of the health and safety risks faced by this population, beyond workers' compensation or other occupational health data sources which do not always capture the full picture of farmworker health issues (18).

Migrant Health Centers and Meaningful Use

As with other health care settings, eligible providers (doctors, nurse practitioners, physician assistants, etc.) in Community and Migrant Health Centers are encouraged to adapt EHR systems and can receive federal monetary incentives if they can show "meaningful use" of the systems in designated periods of time (19). According to the Centers for Medicare and

Medicaid Services (CMS), an EHR is "an electronic version of a patient's medical history, that is maintained by the provider over time, and typically includes patient demographics, progress notes, problems, medications, vital signs, past medical history, family history, immunizations, laboratory data and radiology reports" (20). Since 2011, CMS has offered significant incentives to providers and eligible hospitals that can demonstrate "meaningful use" of EHRs from 2011 to 2016 (19). CMS explains meaningful use in terms of the "The American Recovery and Reinvestment Act (ARRA) of 2009" which stipulates that the certified EHR should 1) be used in a meaningful manner, 2) be able to exchange health information to improve quality of health care and, 3) utilize clinical quality and other measures. Essentially, "meaningful use" is an indicator of quality and quantity that demonstrates EHR technology is used to improve health care. The first two stages of achieving meaningful use involve adapting the technology. The focus of the third stage will be "on achieving improvements in quality, safety and efficiency, focusing on decision support for national high priority conditions, patient access to self management tools, access to comprehensive patient data, and **improving population health outcomes**" (21). Adapting EHRs, making use of HIEs and achieving meaningful use allows eligible providers to improve the care offered at MHCs, obtain financial support and avoid future reimbursement reductions from CMS.

Upcoming EHR requirements regarding meaningful use may help bridge the gap between medicine and epidemiologic surveillance. Rather than conducting expensive studies in patient populations, epidemiologists may be able to utilize EHR and HIE databases to

establish exposure and health outcome associations while addressing meaning use stage 3 requirements, including to "improve population health." The Institute of Medicine (IOM) has charged occupational epidemiologists and other researchers with the task of assessing the feasibility and utility of incorporating occupation and industry variables directly into EHRs; which are not currently required for federally certified EHRs. The IOM makes recommendations regarding coding industry and occupation from text descriptions within the EHR (22) and to utilize such EHR data in population-based studies (13). Systematically incorporating industry and occupation into EHRs can streamline public health research for vulnerable populations as outlined in the third stage of meaningful use.

Public Health Significance

There are many positive initiatives in place to improve the health of farmworkers, including MHCs. Given the availability of clinical data from clinics that serve farmworkers, this may be an efficient way to study specific injuries and illnesses related to farmworkers as compared to other patients served at these clinics. Further, these clinics offer a unique view of farmworkers from an occupational health perspective if coupled with technical, clinical and information support from non-profits such as the National Center for Farmworker Health (23) and the Migrant Clinicians Network (24). Finally, current recommendations from IOM suggest coding industry and occupation directly in EHRs, or retrospectively in a research database, is feasible and can add to the literature of occupational epidemiology. MHCs allow MSFWs to access healthcare and may be vital to understanding the health issues of this vulnerable population. Clinical data available from MHC providers may contribute to the

national discussion on demonstrating meaningful use by the routine inclusion of occupation and industry within certified EMRs (13).

Study Goals

This dissertation analyzed a database assembled from the electronic health records of a Colorado Clinic that serves as a Migrant Health Center (25) with the following goals:

1. To examine the feasibility of using an EHR database to describe the demographics, medical concerns, risk factors associated with hypertension diagnosis, and pesticide exposure of MSFWs compared to other patients at a 2012 Colorado Federally Qualified Health Center (FQHC).
2. To determine if coding of patient occupation and industry within an EHR text field is possible and informative by assessing the percent of successfully coded records.

Specific Aims

The following specific aims addressed the aforementioned hypotheses by utilizing data acquired through the Community-Academic Linkages to Build a National Network for Farmworker Research (Grant# 1RC4ES019405-01)(25). By analyzing a merged, de-identified dataset from a Colorado MHC with adult patient encounters in 2012 across several providers and sites, this dissertation:

1. Described the patients that visited a Colorado FQHC in 2012 by Migrant/Seasonal Farmworker status, including demographic variables, medical indicators and the top medical diagnoses.

2. Assessed the feasibility of coding the text fields that identify a patient's occupation and industry of employment using the NIOSH Industry and Occupation Computerized Coding System (22) compared to manually assigning codes.
3. Described leading risk factors associated with a hypertension diagnosis or elevated blood pressure including MSFW status.
4. Described pesticide exposure where available from nested laboratory values of serum cholinesterase for a subgroup of MSFW patients.

BACKGROUND: REVIEW OF LITERATURE

Health records

In the U.S. alone, medical diagnosis, treatment and intervention represent a multi-billion dollar industry and encompass a large proportion of our gross national product. Surprisingly, this industry is one of the last to adopt electronic systems to track its customers (the patients) and their consumer data (26,27). The medical industry is structured around practitioners- the medical doctors, nurse practitioners, physician assistants and other "providers" of medical care. Each provider contributes to the health record in order to care for the patient, who seeks attention for a health problem (26). Such health problems are categorized into diagnoses, usually coded from the International Classification of Disease (ICD) which is modified in the U.S. to be more supportive of our billing system as the ICD-Clinical Modification (CM) (28).

Over the last century, the health record remained constructed on paper, within folders, in an inconsistent and non-standardized manner. Only recently are health records found in an electronic format. However, billing procedures have been accomplished through postal mail, fax, and also electronic submission, years before the health record evolved to this method. Each provider, or group of providers, implements and sustains their own systematic way of constructing and maintaining such records within patient folders, which do not always match a set of billing codes. The format of the records and the billing codes is different depending on the type of provider and their preferences. If changing providers or needing a specialist, the patient must request his health record be transferred to the new provider or the patient

must recount his medical history to the new provider. Given the variability of health records, it is often difficult to adapt an effective electronic system. With the national adoption of EHRs, a provider should be able to more easily access the records of many of their patients at once, generate reports about his patients, transfer patient information more easily and securely when requested, and also submit billing and prescription requests with better accuracy (26). An EHR is not merely an electronic copy of the patient's paper health record; it is a system that improves the use of the health data that are within the health record. Meaningful use is a measure of quality and quantity that indicates using EHR technology improves health care delivery (21). Most EHRs are built around the international Health Level 7 (HL7), which provides a consistent, comprehensive framework for exchanging, integrating, sharing, and retrieval of electronic health information to support clinical practice and the management, delivery and evaluation of health services including HIEs (29). According to the Office of the National Coordinator for Health (ONC), the EHR must follow this framework and facilitate interoperability between providers to become a federally certified EHR (30).

EHRs have been around for more than 40 years. Several other countries, mainly in Europe, have been making use of network technologies for many years to achieve sophisticated EHR systems (27). Adopting EHRs has been a long process for US providers, fueled mainly by the Health Information Technology for Economic and Clinical Health (HITECH) Act of 2009. The Act provides incentives for eligible providers that can show "meaningful use" of their certified EHR system and will later reduce reimbursement for

providers that do not. A great deal of monetary and policy efforts have been invested into the adoption of the EHR technology and connecting providers through HIE.

Regardless of the late adoption compared to other countries, EHR data poses a great opportunity for many researchers across various fields in America to increase their understanding of disease, injury, surveillance and potential prevention methods (8). Providers and patients in America generally support adoption of HIE (31-33), including among economically disadvantaged groups (34). Public health researchers have already begun using EHR data for surveillance (35-38) and population studies of infectious disease (8). Further, HIEs have also shown promise in improving patient care and reducing superfluous cost by connecting providers to share patient information in a secure manner (9,39,40). Some researchers even found reduced work burden on small providers that use EHRs (41). Patients are now able to access their own health information using "patient portals" through secure internet access (42). The next steps will involve refining the EHR to include important public health indicators, such as the consistent, systematic inclusion of patient occupation and industry as a potential injury or illness risk that can help improve patient diagnosis and treatment.

Occupation information in Electronic Health Records

In 2011, the Institute of Medicine (IOM) released several detailed recommendations on incorporating occupation and industry information into EHRs with meaningful use in mind (13). An ad hoc Committee on Occupational Information and EHRs examined the rationale and feasibility and developed ten recommendations, and five potential benefits. The

ten recommendations include topics such as health information exchanges, meaningful use measurement development, coding and storage challenges, ethics, development, testing, standardization, clinical support, "return-to-work" tools and many other relevant issues that are arising for general health record variables and are anticipated for occupation and industry variables. A detailed listing of the recommendations directly from the IOM report can be found in Appendix A. The five potential benefits of implementing these recommendations are framed within the outcomes and policy priorities of the first stage of meaningful use. These benefits include, 1) improving the "quality, safety, and efficiency of care and reduce health disparities", 2) engaging patients and families in their health care, 3) improving care coordination, 4) improving public health, 5) ensuring protections for personal health information. The committee further concluded that including occupation and industry information into health records may contribute to meaningful use, specifically in relation to population health focus of stage 3 "through better-informed diagnoses, more focused treatment plans, and improved and streamlined return-to-work guidance" (13).

The current peer-reviewed literature regarding EHRs and the value of reporting detailed occupation information about patients is very sparse but increasing. Recent reports from several researchers in occupational health (13,43) have concluded that incorporating occupation and industry history information into EHRs has many benefits to clinical and public health practice. Friedman et al. (44) concluded that including such data into the EHR may assist in correcting billing issues and patient care, but may also standardize occupational health surveillance and injury/illness prevention (44). Lurio et al. (2010) suggested placing

alerts within the EHR framework in order to alert the providers to national and state-level notifiable disease reporting (8) which include some occupational illnesses such as pesticide-related illness (45,46). Such alerts could also be put in place to notify the provider of the potential occupational nature of an illness or injury, but as the IOM notes, the feasibility of such alerts must be assessed before incorporating them as an EHR systems (13). Detailed occupation and industry information in a health record may better inform a clinician's diagnosis and/or treatment. Certain vulnerable populations may benefit greatly from such information being included in their health record.

National Research Agenda for Migrant and Seasonal Farmworkers

In the Migrant and Seasonal Agricultural Worker Protection Act, the U.S. Department of Labor defines a migrant farmworker as "an individual who is employed in agricultural employment of a seasonal or other temporary nature, and who is required to be absent overnight from his permanent place of residence" and a seasonal farmworker as "an individual who is employed in agricultural employment of a seasonal or other temporary nature and is not required to be absent overnight from his permanent place of residence (47)." Other sources describe seasonal farmworkers to be those workers spending at least 51% of their employment on a seasonal basis, engaged in "field and orchard agriculture; packing and sorting procedures in food processing; horticulture specialties (including nursery operations, greenhouse activities, and crops grown under cover); and reforestation" (15,23). NIOSH defines vulnerable populations as groups of individuals that have "limited ability to safeguard their own needs and interests or to seek protections or resources" which may include the

young or elderly, women, those with English-speaking limitations, those with high mobility and/or migration, the disabled, those with low socio-economic status, those with documentation status issues (e.g., immigrant work visa), or with a specific ethnicity and/or culture (12). Therefore, several government agencies recognize farmworkers, particularly migrant farmworkers, as a vulnerable population given their estimated high poverty rates, high rates of foreign-born workers (23), and barriers to healthcare access (12). In 2008, the NIOSH National Occupational Research Agenda (NORA) Agriculture, Forestry, and Fishing (AgFF) Industry Sector Council (12) established an agenda specific to AgFF to establish nine goals for occupational safety and health based on "scientific evidence, public testimonies, peer reviews, and personal expertise." The first five of these goals discusses surveillance, vulnerable workers, outreach, partnerships and communications, and agriculture safety and health (12). However, few comprehensive data sources are available to support these goals among MSFW.

Migrant Health Centers and Research Opportunities

According to HRSA, "Grant-Supported Federally Qualified Health Centers are public and private non-profit health care organizations that meet certain criteria under the Medicare and Medicaid Programs (respectively, Sections 1861(aa)(4) and 1905(l)(2)(B) of the Social Security Act and receive funds under the Health Center Program (Section 330 of the Public Health Service Act))." These centers must be governed by at least 51% of the health consumers. One of the four types of health centers are the Migrant Health Centers. HRSA

currently funds 159 Migrant Health Centers to serve the needs of migratory farmworkers in most of the 50 United States (48,49).

About 862,808 MSFWs and their families were served in 2011 alone (17). The Uniform Data System (UDS) tracks a variety of information in aggregate from these encounters, including patient demographics, services provided, staffing, clinical indicators, utilization rates, costs, and revenues in aggregate. UDS data are collected from all FQHC grantees and reported at the grantee, state, and national levels. As of 2011, approximately 65% of FQHCs (including MHCs) had fully functional EHR systems that were available to all providers at all sites of the FQHC (50).

The National Advisory Council on Migrant Health (NACMH) under HRSA is federally mandated to develop a set of recommendations for action to the Secretary of the Department of Health and Human Services to "increase the effectiveness of MHCs in meeting the primary health care needs of migrant and seasonal farmworkers." The council makes recommendations from information received from national, regional, and community partners, and directly from farmworkers (51). The most recent set of recommendations includes a focus on increased technical support, improved healthcare quality measures and interagency communication (51). These recommendations are consistent with and supportive of meaningful use (21) and with the IOM recommendations (13), may be addressed by incorporating more detailed occupation and industry information into patient health records including EMRs and HIEs with the HL7 compatibility across the country in order to increase understanding of health issues faced by farmworkers.

There may be some concern about the validity of using clinical data to represent farmworkers, however some research conducted suggests the data are similarly representative compared to other survey or surveillance methods. Researchers in New York State compared two methods for estimating occupational morbidity for migrant and seasonal farmworkers (52); their findings indicate that chart review produced similar findings as a random sample worker survey. Brower and colleagues (2009) compared work-related injuries in MSFWs at two major northeast agricultural regions. They found that medical record-based surveillance system, with a correction factor, may effectively estimate occupational illness and injury rates of MSFWs (18). Although only about half of the injuries were treated at a local MHC (18), it is still a valuable tool to reach the MSFW population. Such a medical chart review may supplement current survey methods such as the National Agriculture Worker Survey (53) in describing MSFW health outcomes.

Health issues of Farmworkers

Farmworkers, including MSFWs, have historically faced higher injury, illness, barriers to medical care (54) and even proportionate mortality (55) but only limited, outdated literature is available on the comprehensive medical issues in MSFWs (56). Much of the comprehensive literature available is from more than a decade ago (15,56,57). Smaller, more recent studies show higher burden of chronic diseases such as lung disease (58), hypertension, obesity, and cancer (59). Injuries addressed in the recent literature include hand injuries (60), skin injury from pesticides or other sources (61) and musculoskeletal disorders (54). Aging MSFWs endure persistent musculoskeletal injuries with little use of medical

treatment (62). In New York (63), the top five most prevalent disease conditions among migrant and seasonal farmworkers and their families working in the state from 2003 to 2005 were identified as: 1) infections, 2) musculoskeletal problems, 3) respiratory disease, 4) hypertension, and 5) diabetes. The authors of this study concluded that further action should be taken to serve the health needs of MSFWs and to understand their health needs in comparison to other populations (63). Another study of Michigan farmworkers revealed a perceived need by the MSFWs to improve work-related injuries, chronic illness, and access to dental care (64). A California survey discovered scarce injury prevention programs in the workplace of migrant farmworkers and that half of these workers surveyed complained of musculoskeletal pain in various body parts (65). As a predominately Hispanic, mobile population, several challenges arise when understanding the medical profile of MSFWs, but such research is achievable and could benefit from an update (56).

The National Agricultural Worker Survey (NAWS) is a national, annual survey contracted by the Department of Labor, of approximately 1,500 to 4,000 random farmworkers. Since the survey's inception in 1988, over 53,000 workers have been interviewed to better understand their demographics and work characteristics (53). Several publications have documented findings from NAWS (16,59,66,67), however the most recent analysis of data from the NAWS survey has been examined as a presentation at the National Center for Farmworker Health Midwestern Stream forum (68). Trends from the most recent 10 years of data (1999-2009) reveal that the distribution of farmworkers has remained similar across the "streams" of the U.S. and issues of chronic disease, injury and healthcare access

has not significantly improved. Farmworkers overall are increasing in age, but remain largely Hispanic, foreign-born and poor. The most recent survey results suggests that farmworkers have lived in the United State longer, have improved English speaking and literacy, but still have similar healthcare barriers compared to the previous years' surveys (16,68). Some of the limitations of this survey include that it is only conducted in current workers at their workplace, and there is limited regional and local coverage; it does not include clinical outcomes and specific exposures. However, it is the only reliable, long-term source of national data on agricultural workers and periodically collects more detailed information about social and medical outcomes such as a recent mental health section (53,68).

Wang and others describe some of the major injuries sustained by farmworkers from 1999-2004 utilizing the NAWS. From this investigation, Wang describes the most common injuries as "overexertion from lifting (20%), being struck by hand held objects (13%), and falls to a lower level (10%)." Diversity of occupational injuries may indicate that the best intervention would include focusing on improving overall working conditions and increasing enforcement of existing regulations (69). Further evidence from the NAWS suggests that workers with existing health or musculoskeletal complaints had a higher risk of subsequent injury (66).

Another NAWS study revealed that only 23% of farmworkers reported having health insurance. Most of the insured farmworkers (46%) reported that the farmers paid for their health insurance, with the remaining insured from various sources such as the government (19%), a spouse's insurance plan (12%), or "other" (7%). In terms of workers' compensation

about half of workers (48%) reported coverage by workers' compensation for a work-related illness or injury, while 20 percent reported that they would not be covered and 31 percent did not know (16). Due to low income, most healthcare costs would be prohibitive as an out-of-pocket expense. Migrant farmworkers may also face barriers to utilizing available healthcare even if available due to stigmas surrounding international immigration issues. News reports have discussed this issue in relation to clinic staff. A recent article highlighted the challenges of MHC staff in caring for this group. They visit local migrant workers in the fields in order to provide healthcare and to reach out to the workers for preventive care (70). MSFWs have been reported to use over-the-counter measures to deal with mainly musculoskeletal injuries and chemical burns (71,72)

Conducting research in migratory populations is sometimes difficult, but certainly possible (59). Limited education and mental health issues have been cited as potential concerns to studying farmworkers (64,73). However, a 1995 study about conducting research in MSFWs evaluated the feasibility of identifying MSFWs in their home state and following them over time (74). Due to established community partnerships, permanent home stability and the predictability of seasonal migration patterns, the study team was able to maintain contact with the individuals in the study with only 18 of 196 (9.2%) lost to follow-up. An ecological model has been developed from a subset of the NAWS to describe the relationship between certain social, cultural, and economic health of Hispanic migrant farmworker adults (67). This model identifies language, education and economic barriers to accessing healthcare, and may assist in developing efficient research methods.

Hansen and Donohoe (2003) identify some of the top health concerns faced by farmworkers from occupational hazards and socioeconomic status. Further, the authors point to a need for updating current research on farmworkers to identify contemporary issues and understanding health concerns. Although the population still faces numerous barriers to healthcare, the authors recommend stronger public health infrastructure and improvements to MSFW health status (54). The Healthy People 2020 objectives focus on several issues relevant to MSFWs including healthcare access, environmental health, health communication, health information technology, quality of life, occupational safety and health, public health infrastructure and social determinants of health (75). Occupational information from EHRs can supplement current information available from other surveillance methods, such as trauma registries (76)

Similar efforts in different states have acknowledged MSFW health issues. The California Agricultural Worker Health Survey (CAWHS) reveals similar issues among California MSFWs, such as high rates of uninsured (80%) and high prevalence of chronic disease, such as obesity, high blood pressure, hypercholesterolemia, and diabetes mellitus (65,77). As previously noted, among New York farmworkers: infections, musculoskeletal problems, respiratory disease, hypertension, and diabetes were the five most prevalent diseases identified (63) Another study about New York MSFWs suggests chart reviews with correction factors as a cost-effective estimate of the occupational illness and injury rate in MSFWs (52). Social isolation and working conditions were related to anxiety and depressive symptoms in farmworkers in a sample of 125 male migrant farmworkers in North Carolina

(78). A group of MSFWs in Michigan identified dental health as the most commonly requested service (64) and similar concern was found in California (57). Several other groups have identified needs in MSFWs, but no nationwide system exists to guide the needs of the MSFWs as a whole. EHRs may assist in identifying these needs, in caring for these needs and may already contain occupational information that can help identify MSFWs.

Assessment of Occupation and Industry in Medical Records

"Occupation" can be broadly defined as the type of work an employee performed at a place of work and "industry" as the usual business of the employer (79). Several systematic coding schemes exist for assigning occupation and industry to numeric codes. Occupation and industry are typically coded as two separate variables. The Standard Occupation Classification (SOC) and the 2007 North American Industry Classification System (NAICS) are two common coding schemes for classifying these two variables. SOC codes are a four digit hierarchical coding scheme describing occupation. Similarly, NAICS codes are six-digit hierarchical classification system based on 20 industry schemes and provide industry description and establishment type. The US Census makes use of 3 or 4 digits coding for both industry and occupation. All of these coding systems offer excellent classification of industry and occupation (79).

Coding of industry and occupation available in medical records is of national interest (13) and has been conducted in recent studies to examine the relationship between work and health outcomes (44,80,81). Currently, some clinical settings do record industry and/or occupation information but few utilize a systematic classification system such as the U.S.

Census (82). The fourth IOM recommendation emphasizes NIOSH's high priority on assessing the feasibility of coding available industry and occupation information of patients' EHR even if the information exists in narrative or basic text form (13). Ideally, the industry and occupation will be coded using a specific, updated coding scheme during data entry into the EHR. However, when such entries are simple input text fields, retroactive coding of an EHR database may improve the possibility of occupational epidemiologic research (13).

Extensive time investment can be a barrier to coding industry and occupation within a medical record or in a medical records database. The process involves training of coding personnel in one or more coding schemes and methods which are often updated decennially. In response, NIOSH is currently developing software to improve the coding process and reduce burden on surveillance for industry and occupation including state cancer registries and national surveys. NIOSH released a beta version of industry and occupation coding software in 1998. Although currently outdated, this software was followed by a web-based program that was released to the public in December 2012 (22). This new software, the NIOSH Industry and Occupation Computerized Coding System (NIOCCS), is currently being beta tested by relevant clinicians, health center facilities, cancer registries and vital statistics staff across the nation in order to better understand and improve the functionality of the coding software (22). NIOCCS assigns codes to records using Census 2000 and Census 2002 codes for both industry and occupation, and will also be adding 2010 Census codes. There is also a function to "crosswalk" the Census codes to the NAICS and SOC coding schemes for industry and occupation respectively. Census codes were designed using NAICS

and SOC coding schemes. There are 265 industry codes arranged into 20 sectors, including mining, utilities, wholesale trade, etc. For occupation, there are 509 codes with 23 major groups including management, legal, protective service, construction and extraction, production, etc. The NIOCCS software will be useful to facilitate coding industry and occupation information in an EHR database and reduce the necessary time to code the variables. The software will also streamline the coding of large databases such as EHR databases for epidemiologic research. This study will establish the feasibility of coding such a large database utilizing this software and will compare a common health risk factor (hypertension) among MSFWs and the common industries and occupations.

Hypertension

Hypertension is a global and national health concern. An estimated 30.4% (66.9 million) of Americans are afflicted with this condition which is a risk factor for heart disease and other metabolic conditions. More than half of these (53.5%) are unaware they have hypertension (83). High blood pressure has been identified as one of the most important (83,84) modifiable risk factors for heart disease and stroke: two of the top leading causes of death in the U.S. (85). Even moderate elevation of arterial blood pressure is associated with a shortened life expectancy with increasing epidemiologic evidence suggesting the serious nature of hypertension. Although diet and other lifestyle changes may reduce blood pressure and decrease the risk of associated health complications, medication is often necessary to maintain reduced blood pressure. Inadequate control of high blood pressure can lead to a range of complications, including end-stage renal disease, heart failure and eventual death

(86). Although research about hypertension in MSFWs is limited, hypertension has been identified as a top condition or health concern in several studies (63,64,87) and warrants further investigation to inform preventive and clinical care decisions in MSFWs.

From the 2005-2006 National Health and Nutritional Examination Survey (NHANES), nearly 29% of all U.S. adults of both genders were stage 1 hypertensive or higher (88). A recent study of the NHANES from 1999-2008 and the Behavioral Risk Factor Surveillance System (BRFSS) from 1997-2009 estimated even higher prevalence of hypertension in adults aged 30 years and over. This study estimated hypertension prevalence of 37.6% (range: 26.5 to 54.4%) in men and 40.1 % (range: 28.5 to 57.9%) in women (89). However, this prevalence estimate is different for minorities, and higher in African-Americans (89,90) although it is increasing in all race/ethnicity groups (91). Researchers from the Framingham study (92) found that hypertension rarely occurs in isolation. In about 80% of cases, another condition is present, most commonly glucose intolerance, obesity, left ventricular hypertrophy, and dislipidemia (elevated total, LDL, and small-dense LDL cholesterol levels, raised triglyceride, and reduced HDL cholesterol levels). Additionally, study results suggest that clusters of three or more of these additional conditions occur at four times the rate expected by chance. Alarmingly, the risk of coronary disease (often lethal) increases stepwise risk factor clustering (92).

Currently, limited information exists regarding the risk factors associated with hypertension for MSFWs. There is evidence that Hispanics may be at higher risk of hypertension (93) and the majority of MSFWs are Hispanic, predominately Mexican

(15,16,68). A recent study showed that Texas MSFWs have higher proportionate mortality due to hypertension among other risk factors (55). Also, national data indicates Hispanics may be at higher risk for complications from untreated hypertension compared to other workers of similar age, gender, education and income (83,93). As a predominately Hispanic population, further studies are necessary to better understand the prevalence and risk factors of hypertension in MSFWs.

According to Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure (JNC 7) (86), hypertension, or high blood pressure, is defined as follows in adults (18 years or older) In the clinical setting, the classifications are "based on the average of two or more properly measured, seated blood pressure readings on each of two or more office visits."(86):

- Normal—systolic blood pressure (SBP) is lower than 120 mm Hg; **and** diastolic blood pressure (DPB) is lower than 80 mm Hg
- Pre-hypertension—SBP is 120 to139 mm Hg; **or** DBP is 80 to 89 mm Hg
- Stage 1—SBP is 140 to159 mm Hg; **or** DBP is 90 to 99 mm Hg
- Stage 2—SBP is equal to or more than 160 mm Hg; **or** DBP is equal to or more than 100 mm Hg

Although this definition may seem straightforward, variation in the measurement of blood pressure may pose issues of validity (94). For the purpose of this study a diagnosis of hypertension (ICD-9CM codes 401.X-405.X) was used in addition to blood pressure values

at the last clinic visit. To address specific aim 3, a detailed analysis will be conducted in order to compare the prevalence of elevated blood pressure (stage 1 or stage 2, identified by the ICD-9CM codes 401.X-405.X)

Work-related Pesticide Exposure

A recent analysis of nationally reported acute-pesticide poisoning (APP) cases, revealed that 71% of patients were employed as farmworkers (95). The results of this study show that APP is still a health issue disproportionately affecting farmworkers. Additionally, the risk of APP in MSFWs varies by many factors including gender, age location and crop but have remained high over the study period from 1998 to 2005 (95). APP is typically not life-threatening, but can have a wide range of symptoms and health consequences including death. Although education about safe pesticide practice has increased in agricultural workers, workers have reported handling pesticides more often compared to previous years (68).

A pesticide is "any substance or mixture of substances intended for preventing, destroying, repelling, or mitigating pests. Pests can be insects, rodents, weeds, and a host of other unwanted organisms (96)." Pesticides are typically classified by their target such as insecticides, herbicides, fungicides, and rodenticides. Advanced classification includes species-specific pesticides, plant growth regulators, repellants, and attractants. Over the last 60 years, many detailed developments in the chemistry, nomenclature, biotransformation and degradation, environmental effects and toxicity of pesticides have accompanied many new pesticides. Some of the most commonly used active ingredients in the American agriculture industry include many herbicides and insecticides such organophosphates. As of 2000, an

estimated 75% of households in America use some form pesticide to ward off unwanted houseguests. Globally, pesticide production is a multi-billion dollar industry (96). According to a recent report from the Environmental Protection Agency (EPA), the agriculture industry sector in America spends over $7 billion on pesticides annually. This represents approximately two-thirds of the total expenditure making agriculture the largest buyer of pesticides in America (97).

In the human body, cholinesterase limits acetylcholine in the nervous system at the neuromuscular junction to control movements of the muscles. If interrupted by a cholinesterase inhibitor (such as certain chemicals containing organophosphates (OPs) and carbamates (CMs)) acetylcholine begins to accumulate in the blood and the nervous system cannot control movements of the muscles (98). Common symptoms of APP, therefore, are neurological, such as muscle twitching, vision impairment, eye pain, constricted pupils, respiratory symptoms, diaphoresis, seizures, etc. The major signs and symptoms include overstimulation of exocrine glands, smooth muscle stimulation, gastrointestinal symptoms, hypertension, overstimulation of skeletal muscle with subsequent fatigue and central nervous effects (98).

Pesticides are unique in that humans deliberately use their toxic effects to kill living organisms, but pesticides do not always cause injury to pests alone (96). One of the most common poisonings worldwide results from exposure to organophosphate pesticides. Approximately one million cases of such poisoning and several hundred thousands of deaths occur annually around the world. Depending on the type of pesticide, route of exposure, the

age, gender, physical health of the victim and many other factors, clinical symptoms are variable and often difficult to diagnose (99). As a nationally reportable disease, the definition of APP is designed to be flexible. A case is defined by an acute onset of symptoms that are related to the pesticide product formation and involve one or more of the following symptoms: “Systemic signs or symptoms (including respiratory, gastrointestinal, allergic and neurological signs/symptoms); Dermatologic lesions; and/or Ocular lesions (100).” A common marker for pesticide exposure, specifically organophosphates, is the reduction of serum cholinesterase (100). In addition to serum analyses, several methods have been used to ascertain pesticide exposure including dust wipes (101), urinalysis (102), but assessment methods should be tailored to the type of pesticide exposure. Exposure to pesticides can occur through dermal absorption, inhalation or ingestion. Those at highest risk of exposure include workers that produce, transport, mix, load or apply pesticides, with dermal exposure thought to be the most important route of exposure although not always most hazardous (96). Pesticide-related illness and injury is a nationally reportable disease but only certain states are funded to report such cases (46).

Several small-scale studies of pesticide exposures have been conducted to assess the effects of APP from pesticides containing OPs and CMs. OPs are well known to cause progressive neurological damage from cumulative exposures (99). Neuroprotective therapies exist to prevent brain damage resulting from cumulative irreversible AChE inhibition when recognized by a clinician (99).

Despite many federal and state programs aimed at reducing APP, agricultural workers, including MSFWs and farmers, are at high risk of exposure to pesticides. The EPA developed the Worker Protection Standard for Agricultural Pesticides (WPS) as a regulation to "reduce the risk of pesticide poisonings and injuries among agricultural workers and pesticide handlers." This regulation requires workers to be away from areas treated with pesticide; personal protective equipment (PPE) to be provided and maintained for handlers and workers while in a treated area; decontamination supplies and emergency assistance for accidental poisoning or injury; and workers must also be informed and trained on pesticide composition and handling procedures (103). Although several exceptions exist to this standard, the idea is to reduce exposures of agriculture workers to potentially toxic chemicals. The WPS requires this safety training be provided by employers, but regulations regarding pesticide exposure protection vary by state (104). NIOSH has also established the Sentinel Event Notification System for Occupational Risks - Pesticides Program (SENSOR-Pesticides) to "reduce the number of injuries and illnesses associated with occupational pesticide exposure (100)." Physicians, other health professionals, and laboratories are required to report pesticide exposures to state health departments that participate in the SENSOR program, such as Texas (105), California and nine other states. .

A recent California surveillance study identified acute pesticide-related illness and injury cases among farmworkers using SENSOR. The researchers found an increased risk for acute pesticide-related illness and injury among female farmworkers, but the absolute number of such cases is higher among males than females. This may be due to physiological

differences between males and females (106). The researchers also recommend improved regulation enforcement to protect against off-target drift exposures to reduce acute dermal and respiratory exposures (106). Further health concerns may arise after chronic exposure to pesticides. Such health concerns exist in exposure to pesticides such as cancer (59) and oxidative stress (107). A recent surveillance case investigation revealed that pesticides also have an impact on reproductive health, including birth defects (108).

Pesticide training has been shown to be effective in increasing safe handling practices. A group of North Carolina MSFWs had significantly higher use of safe procedures with training (109), but the same researchers also point out that not all farmers are providing the necessary tools to reduce exposure (such as housing with bathing and laundry facilities) (110). Results from a NIOSH cohort study in South Texas suggest high variability in those farmworker families receiving training (111,112) with as low as 18.8% receiving training in some areas. In another study about attitudes and practices, the majority of farmworkers (87.5%) acknowledged the potentially harmful effects of pesticides but more than half (55.4%) said that they rarely or never wear PPE such as gloves (113). Informal training among farmworkers may also prove to be helpful in increasing safe procedures (114).

Summary of Literature

Occupational epidemiology is a well-established field of public health that has improved the health and safety of workers across the world. Nationwide adaption of EHR technology may provide innovative sources of data to examine workplace exposure-outcome relationships. MHCs are currently in the process of adopting EHR systems and

predominately serve vulnerable populations including MSFWs. Few national databases are available to study the health and safety and of MSFWs, and a national Health Information Exchange of EHR data from MHCs may provide a unique look at the health of MSFWs and improve the surveillance and occupational epidemiology of health and safety issues of MSFWs. The purpose of this study is to examine the feasibility of using an EHR database in an epidemiologic study of a vulnerable working population.

METHODS

EHR Data Source and Population

The Community Based Research Network (CBRN) is a National Institute of Environmental Health Sciences (NIEHS) funded research project (Grant # 1RC4ES019405) aiming to establish a "National Farmworker Research Network (NFRN) consisting of a national consortium of selected Community/Migrant Health Centers supported by an Academic Health Center and a national representative Community Research Associate" (25). Further aims of this research project include establishing a centralized, computerized, and secure National Farmworker Clinical Research Database (NFCRD). This database consists of medical records with the purpose of providing researchers with either a limited or de-identified medical database devoid of most or all personal identifiers (18 HIPAA identifying variables, or Protected Health Information (PHI), such as name, birth date, address, etc. as outlined in Appendix B). Since 2009, the network has built partnerships starting with the multiple principal investigators representing the farmworker community around the U.S., including academic, clinical, and community partners for the purpose of infrastructure building. This database is owned and shared by the participating clinics where data sharing will be decided by the CBRN steering committee using specified guidelines and agreements (25). The best method to form this database was to merge all of the participating clinics into a Health Information Exchange (HIE). Centex Systems Support Services (CSSS) is a CBRN partner and serves as the expert in clinical data HIE, database manager and extension of the clinics through a Business Associate Agreement. After Institutional Review Board approval

from the University of Texas Health Science Center Committee for the Protection of Human Subject, and CBRN steering committee approval, a de-identified clinical dataset, was transferred for analysis from the database managers (CSSS).

The first clinic to become part of the HIE was the Salud Family Health Centers of Colorado. Operating as nine community health clinics and a mobile unit, Salud prioritizes service to all low-income, medically underserved, migrant and seasonal farmworker community members (115).

Study Design

This cross-sectional study captured a snapshot of de-identified patient EHR data from January 1st through December 31st in 2012 from all clinic sites of Salud. The database included patient demographics, vital sign measurements from clinic visits, a free text variable with some occupation and/or industry information, as well as a subset of 61 serum cholinesterase lab values.

Human Subjects Considerations

This study analyzed an EHR database, just as though the records were abstracted from paper health records. Patients were receiving their usual medical care at the clinics. No participation was required on the part of the patients since a previous IRB-approved Waiver of Authorization was established for release of records rather than an individual informed consent; no specimen collection or clinical measurement took place, and this research consisted of data abstraction on a large scale. All PHI, as outlined in Appendix B, was maintained separate from the research database and was inaccessible to the investigator. A

data use agreement (Appendix C) was agreed upon by the three principal investigators (PIs) of the CBRN project to outline specific data handling responsibilities and intended use of the data by the dissertation researcher. The physical clinic from which the patient data originated was identified by a random numeric identification code. All patients included in this study were adults, aged 18 years and older, and any identifying information was removed before receipt by the investigator.

Data Acquisition

Data were acquired using a Virtual Private Network (VPN) to securely transfer data from the clinic to the HIE manager as a pilot. VPNs are secured connections that allow encrypted transfer of data and will also serve as a tunnel through which the HIE manager can regularly pull data from the clinic into the HIE for future updates. There were some issues establishing this connection using the HL7 coding interoperability standards, so this initial exchange via VPN pulled the data using SQL coding. CSSS used the information gathered to interpret the EHR data back into HL7 (29). The data are stored in specific, secure warehousing units and accessed by approved CSSS analysts only for database building and reporting. CSSS received the data and removed any PHI as outlined in Appendix B, and encrypted patient ID. CSSS securely transferred this dataset to the investigator using password protected, University of Texas secure cloud storage (X-files).

Description of Patients (Specific Aim 1)

Records were reviewed for errors and completeness. Upon review, the format of the data were found to be grouped by encrypted encounter identification numbers, so it was necessary to remove duplication of patient data and regroup by (encrypted) patient identification numbers (IDs). Further, encrypted patient ID combination number/text records were replaced with shorter digit values to meet the variable formatting requirement specified by the data analysis software. Demographic factors available from the patient record included: age, marital status, sex, primary language, race, ethnicity, insurance, status of the following: public housing, homeless, employment, migrant farmworker, and seasonal farmworker. Migrant and Seasonal farmworkers were grouped together as "MSFW". Medical indicators from the patient encounter included: body mass index (BMI), temperature, heart rate (pulse), and systolic and diastolic blood pressure. BMI was calculated in the medical record, provided in the dataset, and edited for out-of-range values. Additionally, a binary variable was generated to indicate patients with a BMI greater than or equal to 30 kg/m^2 to reflect obesity versus other. After de-duplication, the last value of the medical indicator in 2012 was included in the study (i.e. that most recent encounter's BMI value). If multiple values existed for the last encounter, an average was taken. Diagnosis codes were also de-duplicated and counted by illness, excluding codes for preventive/non-illness procedures such as vaccination or pregnancy. From the diagnosis codes, binary variables were created to indicate patients with a 2012 diagnosis of one of the most frequent counted diagnoses for all patients including hypertension, overweight/obesity, type 2 diabetes, lipid disorder, back

disorder, anxiety, joint disorders, abdominal pain, drug abuse, and soft tissue disorders. A binary variable was also generated for patients with a history of any surgical procedures. The following variables were excluded due to text entry and a lack of systematic coding scheme: allergies, education, income, family history, medication list, and social history (drug use, alcohol use, sexual orientation).

The patient population is described as a whole and by MSFW status in Table 1 regarding demographics. Since a large portion of patient records were missing MSFW status information, we categorized patients with null values for MSFW status as a third comparison group. Medical indicators were summarized including the 10 most frequent clinical diagnoses based on ICD-9CM code counts for all patients and by MSFW status in Table 2. The chi-square test or the Fisher's exact test was used to examine differences for categorical variables (e.g. sex, race, etc.) between MSFWs versus non-MSFW patients and patients with no information (No Info) regarding MSFW status. The Bonferroni adjustment was used to adjust the resulting P-value for pairwise multiple comparisons for the three MSFW groups (MSFW vs. Non-MSFW, MSFW vs. No Info, and Non-MSFW vs. No Info) to make sure that the overall Type I error rate was controlled at 0.05. For continuous variables (e.g., blood pressure, BMI, etc.), the Kruskal-Wallis non-parametric test was used to assess the equality of the median across the three MSFW groups. Further, the Dunn procedure was used to conduct pairwise multiple comparisons of MSFWs, non-MSFW, and no info groups and Bonferroni adjusted p values were computed for each pairwise comparison. All statistical

tests were conducted at a significance level of 0.05 (two-sided). Statistical analyses were conducted using Stata version 11.0 (StataCorp, College Station, TX (116)).

Feasibility of Assigning Industry and Occupation Codes (Specific Aim 2)

The NIOSH Industry and Occupation Computerized Coding System (NIOCCS) was released for public use in December 2012. This software automatically codes text field occupation and industry variables at three different levels of accuracy including 30%, 70%, and 90% confidence according to the 2000 or 2002 U.S. Census coding index for Industry and Occupation (22). The lowest percent confidence has the highest rate of error, but yields the highest percentage of coded records (i.e. the least missing data). The highest percent confidence codes text fields with the most accuracy but yields the lowest percent of successfully coded records. A blank is different than an "unknown". When an industry or occupation is clearly reported as "unknown", "not currently working", "not recorded", etc., the software or manual coder assigned "999" for industry and "990" for occupation. However, if the software is unable to understand the text available in the field (such as an unfamiliar business name or a gross misspelling), the software leaves that code blank and indicate the need for manual coding.

Manual coding served as a gold standard and required searching through the Census Industry and Occupation Classification Manual and Index (117). The NIOCCS software automatically codes multiple records according to the desired coding scheme (Census 2000 or 2002). The investigator was trained by NIOSH on the 2000 Census coding scheme (82) and the NIOCCS coding software (22). As mentioned above, a manual coder leaves no blank

records and if the industry or occupation cannot be understood from the text (or the text field is blank), the manual coder assigned “999” to industry and/or “990” to occupation. A manual coder, however, is able to perform an internet search for an unfamiliar business name or recognize the text as misspelled. If the manual coder cannot understand the text, then the final decision to code as 999 for industry or 990 for occupation is made. The manual coder rarely leaves a blank because they can use the 999/990 codes. The amount of time to successfully code 1000 records manually was estimated to be one week of full-time employment by NIOCCS staff (FTE, approximately 40 hours, personal communication NIOSH, 10/4/2012).

The automatic coding process is outlined in detail in a manual on the NIOCCS website (22). Records were formatted in a tab delimited .txt file from a spreadsheet software and included the following variables as required by the software, : a unique identifier, industry text, occupation text, employer company name, job duties, employer city, employer state, employer zip, age, education level, user defined field 1, user defined field 2. The user defined fields do not require any specified format and for the purpose of this study MSFW status and employment status were included. Files uploaded must be less than one megabyte in size or the software will yield an error, therefore the files were split into smaller files and uploaded for coding at all three confidence levels. A text entry for occupation was available for 33,868 patients in 2012, as displayed in Table 3. The non-blank occupation descriptions varied highly in their content with some entries detailing more than simply occupation information. Therefore, the entries for “occupation” were included under the corresponding

variable heading for “industry”, “occupation”, “employer company name”, and “job duties” for submission to the software. The NIOCCS software does not assign a code for unrecognized text entries, but does assign the “unknown” code (999 for industry and 990 for occupation) for text entries that clearly say “unknown”, or “not currently working”, etc. If a text entry is the name of a business, for example, the software will typically be unable to assign a code and will leave that record blank.

The percentage of successfully coded records by accuracy level is shown in Table 3. All records were assigned a hash code for patient identification encryption and randomized based on the patient ID. A simple random sample of 100 records was manually coded using the methods from above and compared to the automatic coding. The degree of agreement between manual coding and software auto coding at different accuracy levels was examined using Cohen's kappa statistic for industry and occupation outcomes separately. Cohen’s Kappa is a measure of agreement that takes into account the amount of agreement which would be expected by chance, and it will be interpreted as following: less than zero is poor agreement, from 0-0.20 slight agreement, from 0.21-0.40 fair agreement, from 0.41-0.60 moderate agreement, from 0.61-0.80 substantial agreement, and from 0.81-1 almost perfect agreement (118) In addition, the time required for manual and auto-coding the EHR database was recorded to serve as a comparison with this estimation for future feasibility estimation.

Analysis of Prevalence of Hypertension (Specific Aim 3)

Hypertension is a chronic health problem associated with heart disease (85,86) and as previously stated, has also been identified as a top health concern in MSFWs (63). A large EHR database provides an opportunity to examine potential risk factors associated with hypertension in both MSFW and their peers. All data were collected in a clinical setting by providers at a Colorado clinic with multiple sites. To address specific aim 3, we first examined the bivariate association between the presence of a hypertension diagnosis in 2012 (ICD-9CM codes 401.X-405.X) and the MSFW status without adjusting for confounding variables Table 4. A multivariable logistic regression analysis was then used in order to assess potential risk factors of clinical diagnosis of hypertension (ICD-9CM codes 401.X-405.X). Two full logistic regression models were generated using all available demographic, clinical variables, binary diagnosis variables, and MSFW status. The first model used hypertension diagnosis as the main outcome of interest and the second model used elevated blood pressure as defined by the JNC 7 (86). A backward model selection method was used to remove one variable at a time from each model, resulting in a final model with either statistically significant ($P<0.10$) or clinically important variables. Adjusted odds ratios and corresponding confidence intervals were summarized by variable for the final model in Table 4. Interactions of age and obesity, sex and obesity and age and sex were examined.

Pesticide Exposure Pilot Analysis (Specific Aim 4)

At the Salud Family Health Center, 61 serum cholinesterase laboratory assessments were collected on patients suspected to be exposed to organophosphate or carbamate pesticides as part of their routine clinical care (cholinesterase is a known biomarker of organophosphate pesticide exposure (119)). Specific days of farm work of the patient were not collected as a variable and may have been collected within the EHR, but were not transferred as part of the database for this research. Seasons of harvesting or processing vary depending on the fruit or vegetable growing and the nature of the work of the MSFW, and the geographic regions of the crop, so the exact dates of laboratory values varied depending on the patient appointment and follow-up. Generally, the harvest season occurs May through November depending on the crop and the off season occurs during about December through April. The general range of expected serum cholinesterase values is 3342 - 7586 IU/L for males, and 2673 - 6592 IU/L for females. However, in order to interpret a cholinesterase value, it must be compared to a baseline value of the same patient. To collect a baseline value, patients were instructed to reduce or eliminate exposure and re-test approximately a month later to understand whether the exposed value was substantially lower than an unexposed "baseline" value (120). Clinical manifestation of symptoms of pesticide poisoning typically occurs at greater than 19% reduction of cholinesterase activity during exposure compared to baseline. Although some ICD-9CM codes may be indicative of APP (such as E863- Accidental poisoning by ingestion of organophosphorus compounds), no specific code is assigned for APP. Depending on state laws, Workers' Compensation Insurance may cover

pesticide-related medical costs. However, coverage requirements for the agricultural industry vary by state and may depend on the worker's legal status, time of employment and other factors (121). A 25% reduction of serum cholinesterase from baseline (pre-planting) to potential exposure (harvest season) and accompanying symptoms is considered compensable by Workers' Compensation at this specific clinic and is limited to the circumstances surrounding the patient's employment. (Personal Communication Ed Hendrickson, PhD, PA: 11/5/2012). For this study the outcome of interest was the percent reduction in cholinesterase level from the estimated off-season (baseline) to the harvesting season (suspected exposure). Exposure assessment would be difficult to ascertain from the EHR, however, a review of available information from the "problem list" was also discussed in the results.

RESULTS

As outlined in Figure 1, we identified 98,175 patient records from a Colorado FQHC in 2012. Patient records were excluded from the analysis if the patient was less than 18 years old (n=28,979), or if all medical indicators were missing for encounters in 2012, indicating no provider visit during 2012 (n=27,379). As a result, a total of 41,817 patients were included in the study. Of these, 553 were indicated as MSFW patients; 20,665 were identified as non-MSFW patients and 20,599 had no information in the MSFW field. The No-Info group had several variables with mostly missing data compared to the MSFW or Non-MSFW patient groups.

Patient Demographics (Specific Aim 1)

Patient demographics (age, sex, marital status, employment, preferred language, usage of public housing, homelessness, race/ethnicity, and insurance) were described and compared across MSFW status groups in Table 1. A p-value for testing the overall difference among the three groups is indicated in the last column. Pairwise comparisons between MSFWs compared to Non-MSFWs and the No Information groups were conducted with significant differences indicated with an asterisk if a significant difference was observed compared to MSFWs after Bonferroni adjustment for multiple comparisons.

Among the 41,817 study patients, the median age was 39 years, with males comprising 38.5% of the population. 41.2% of patients identified as married; 34.5% were designated as employed; and the majority of patients identified English as their primary language (67.0%). Few patients overall were identified as Public Housing Users (2.0%) or

Homeless (3.3%). Most patients identified as Hispanic (53.7%) or non-Hispanic white (40.6%). The majority of patients utilized federal or state funding as payment for medical service (81.7%) compared to self-pay (4.0%) or private insurance (14.3%).

As compared to Non-MSFWs, MSFWs were significantly younger, had a higher proportion of males and a higher percentage of patients identified as married. MSFWs were identified as employed more often than Non-MSFWs and identified their primary language as English significantly less often. MSFWs were also significantly less likely to identify as homeless, but had a similarly small percentage of public housing users. The majority of MSFWs were Hispanic (85.8%) compared to about half of Non-MSFWs (49.1%). MSFWs were more likely to use self-pay rather than the federal or state funding to pay for their care compared to the Non-MSFW group.

As compared to the No-Info group, MSFWs were significantly older, and had a significantly higher proportion of male patients. MSFWs had a significantly higher percentage of married patients, a somewhat higher employed percentage, and a significantly lower percentage of primary English speakers. MSFWs had a significantly higher percentage of Hispanic patients with only half (50.9%) of MSFWs utilizing federal/state funding for medical compared to 72.7% of the No-Info group using federal/state funding. The No-Info group had a very high number of missing values in the marital status, employment, language, public housing, homeless and insurance variables, which may lead to bias and unreliable results when comparing the No-Info group to MSFWs.

Although not displayed in a table, we also compared the Non-MSFW group to the No-Info group. Compared to the Non-MSFW group, the No-Info group was significantly younger (P<0.001), and had significantly more Hispanic patients (P<0.001). Other comparisons were not significantly different.

Patient Medical indicators (Specific Aim 1)

Table 2a displays the 10 most frequent patient medical indicators and the prevalence of the 10 most frequent diagnoses in Table 2b.

BMI and Obesity: The mean BMI for all patients at the last visit in 2012 was 29.3 kg/m^2, with 39.3% classified as obese. Overall, the MSFW group had a lower mean BMI compared to the Non-MSFW group but a significantly higher mean BMI compared to No-Info group. The percentage of obesity was significantly higher in the MSFW and Non-MSFW groups compared to the No-Info group.

Blood Pressure (BP) and BP Categories: Median blood pressure was 120/78mmHg for all patients at the last visit. The MSFW group had significantly higher median systolic BP compared to Non-MSFW and No-Info group patients. MSFW patients had significantly lower diastolic BP compared to Non-MSFW but significantly higher compared to No-Info. When BP was categorized according to normal, pre-hypertension, hypertension stage 1 and hypertension stage 2 (82), more than half of all patients were categorized as pre-hypertensive (43.4%) or hypertensive (20.1%). The percentage of normal BPs was similar between MSFW and the Non-MSFW patients but the No-Info group had a significantly higher percentage of patients with a normal BP.

Vitals: Median heart rate was 80 beats per minute and mean patient temperature was 98.03ºF. Overall, 66.1% of patients had a previous surgical history. MSFWs had significantly lower heart rate, lower percentage of past surgery and no difference in temperature compared to Non-MSFW or No-Info patients.

Diagnoses: each patient had a list of diagnoses associated with 2012 clinic visits. A patient may have had more than one diagnosis for a given visit, so we removed any duplicates. Diagnostic/preventive (V-codes) or Unknown (799) codes were excluded. The most frequent diagnosis codes for all patients were displayed in Table 2b. The composition of the most frequent diagnoses in all groups was similar, but the leading diagnosis in MSFWs was overweight/obesity followed by hypertension. The prevalence of each of the 10 most frequent indicators was lower in the No-Info group except for back disorder diagnoses that were the leading diagnosis (15.6%). Note that diagnoses were not mutually exclusive; rather, it was possible for a patient to have multiple diagnoses. Therefore, prevalence percentages across diagnoses rows cannot be added.

Industry and Occupation Coding (Specific Aim 2)

Of the 41,817 patient records included in the study, 33,868 (81.0%) non-blank occupation text field records were submitted to the NIOCCS software for coding. Text available in the occupation text field was copied and pasted into the fields for “industry”, “job duties”, and “employer name” according to the NIOCCS spreadsheet template. Coding time was approximately 30 minutes for 100 records of both industry and occupation. Auto-

coding of the 33,868 records took less than 12 hours from submission with no other users in the queue for coding.

In Table 3, the results of the successfully coded records are displayed under the "number of records" and "percent coded". Manual coding of 100 records was conducted and compared to the same 100 records coded by the NIOCCS software at the three different confidences. In Table 3, two sets of Kappa statistics are displayed. The first set displays the Kappa for industry and occupation that included blanks in the analysis. The second set of Kappa excludes blanks from the software to test the accuracy of actual coded records (rather than comparing blanks to manually coded records). Since the software leaves uncoded records blank, this lowered the agreement between the software and manually coded records. Even with the blanks, agreement was significantly higher than would be expected by chance, and fell within the "substantial agreement" category for kappa, except the 90% confidence group for the kappa that included blanks. Contingency tables showing the joint distribution of the two different coding methods are displayed in Appendix D.

Of the 21,212 records successfully coded at 90% confidence, 10,707 (50.5%) were coded as "unknown" industry and occupation (999/990). Such codes are assigned by NIOCCS or a manual coder for text entries of "unknown" or "not currently working," etc. Therefore, 10,505 patient records (25.1% of 41,817) were successfully recorded in the EHR and coded at 90% confidence (excluding unknowns).

Hypertension Risk Factors (Specific Aim 3)

A logistic regression model was constructed to understand the potential risk factors associated with a hypertension diagnosis based on the ICD-9cm codes given in the Methods Section. The full model included all demographic variables available and all medical indicators except blood pressure measurements, pulse and temperature, and the interactions of age and obesity, sex and obesity, and age and sex.

Another regression model was constructed using the medical definition of pre-hypertension or higher (above 120 mmHg systolic **or** 80 mmHg diastolic) as a dichotomous outcome for blood pressure at the latest visit in 2012 for each patient. The full model included all demographic variables available and all medical indicators except for hypertension diagnosis, pulse and temperature. The final, reduced models are displayed in Table 4 for both hypertension diagnosis and elevated blood pressure.

The following factors were significantly associated with being diagnosed with hypertension: being an older man, having a previous surgery history, being obese, or having clinical diagnoses of lipid disorder, type 2 diabetes, anxiety, drug abuse, or soft tissue disorder. In addition, the interaction between obesity and age, and the interaction between sex and age were found to be significant. After adjustment, for every one-year increase in age, the odds of having a diagnosis of hypertension increased by 8.7% (OR=1.087, 95% CI: 1.081, 1.093), 7.6% (OR=1.076, 95% CI: 1.071, 1.082), 6.1% (OR=1.061, 95% CI: 1.055, 1.067), and 5.1% (OR=1.051, 95% CI: 1.044, 1.057) for non-obese women, obese women, non-obese men, and obese men, respectively. Men were more likely to be diagnosed with

hypertension than women, however, the gender difference was reduced with aging (30-year old: OR=2.598, 95% CI: 2.234, 3.021; 40-year old: OR=2.038, 95% CI: 1.836, 2.264; 50-year old: OR=1.599, 95% CI: 1.472, 1.737). Similarly, obese people were more likely to be diagnosed with hypertension than their non-obese counterparts, however, the difference was reduced with aging as well (30-year old: OR=2.665, 95% CI: 2.289, 3.103; 40-year old: OR=2.408, 95% CI: 2.168, 2.675; 50-year old: OR=2.176, 95% CI: 2.004, 2.362). Similar associations were observed for elevated blood pressure. The risk factors that were significantly associated with elevated blood pressure include: being an older man, being an English speaker, being obese, or having clinical diagnoses of lipid disorder or type 2 diabetes. In addition, the interaction between obesity and age, and the interaction between sex and age were found to be significant. MSFW status was not associated with either hypertension diagnosis or elevated blood pressure after adjusting for other factors.

Assessing Exposure to Pesticides (Specific Aim 4)

As part of routine patient care, 61 values were collected from patient laboratory values. Of these, 36 values from 36 different patients were single values with no comparison, and 25 values were from patients with paired values. Displayed in Table 5, 12 patients had both a baseline and exposed cholinesterase value over two separate visits for a total of 24 paired values (one patient had three lab values). All values were in the normal expected range except patient 12. Patient 12 showed no major change between measurements, but remained below normal in both values indicating no “baseline” value was obtained. All other patients showed a decrease in cholinesterase from baseline to suspected exposure indicating some

potential exposure to organophosphate or carbamate pesticide. As mentioned above, the major signs and symptoms include neurological disturbances (twitching, constricted pupils, seizures, etc), overstimulation of exocrine glands, smooth muscle stimulation, gastrointestinal symptoms, overstimulation of skeletal muscle with subsequent fatigue and central nervous effects (98). In the "problem list" section, symptoms related to pesticide exposure were found in six patients without paired cholinesterase values, and included abdominal pain, dizziness, facial droop, fatigue and malaise, headache, hypertension and urinary incontinence. Twenty-three of all patients with cholinesterase values had no information in the problem list section. No potentially related symptoms were noted in the records of patients with paired cholinesterase values. No information was available in the database regarding exposure.

DISCUSSION

No national HIE currently exists in the U.S. As potentially the first anchor for a national HIE, the National Farmworker Clinical Research Database (NFCRD) will address a lack of current, geographically diverse medical data available for farmworkers around the country and encourages progress in HIE across America. The first clinic added to the HIE was a Colorado FQHC that serves as a MHC to nine communities in northeastern Colorado. Patient record data from 2012 were transferred from the health center to CSSS, the HIE manager. Data from this initial transfer were de-identified and transferred securely to the investigator for analysis. Information regarding patient demographics and medical indicators were examined for feasibility of study with results presented above.

Summary of Results

Specific Aim 1: Overall patients were, on average, just under 40 years old; Hispanic or white non-Hispanic; and were primarily female. Fewer than half reported being married, employed, or non-English speaking. Most patients relied on state or federal programs to cover medical expenses. Compared to Non-MSFW patients, MSFWs were younger, had more male patients, more married patients, more employed patients, fewer English speakers, more Hispanics and a higher proportion of self-paying patients. The demographic findings from this study were similar to those found in the NAWS survey, including the high percentage of Hispanics among MSFWs and using self pay to cover medical expenses (16,68). The No-Info group had a significant number of missing records in other variables including marital status, employment, language, public housing status, homeless status, and

insurance (greater than 99% missing). These missing values made it difficult to interpret significant differences between the No-Info group and the other patient groups. The No-Info group is suspected to have a combination of MSFW and Non-MSFW patients, and was therefore still included in the patient description as a separate patient group. Few comprehensive studies exist to characterize demographics of MSFWs nationally, and specifically in Colorado. Future versions of this database should address the missing information with improved experience in HIE technology and will include patient data from several different areas around the country.

In regards to medical indicators, there was a high percentage of obese patients (39.3%), and mean BMI was just under obese classification. Median systolic and diastolic blood pressures were close to the pre-hypertension cutoff, with only 36.6% of patients below the pre-hypertension cutoff at the last visit (120/80 mmHg). Most patients had a normal heart rate and temperature; over 66% of patients reported some form of previous surgery. The most frequent diagnoses assigned to patients in 2012 revealed chronic disease indicators and there were similarly high diagnosis prevalence for MSFWs, Non-MSFWs and No-Info patients. Most differences between patient groups observed in regards to medical indicators were statistically significant, but not always clinically significant. The large sample size and use of non-parametric statistical methods may have resulted in statistical significance more so than clinical significance. Most of the information available regarding the health of MSFWs is specific to certain regions, such as in California (57), Michigan (56,64), New York (63), or Texas(56). Limited national information is available from the NAWS (16,68). In the

medical literature available, type 2 diabetes and hypertension were always among the top diagnoses, as shown in this study (56,63). Future expansion of this analysis to other FQHCs will provide the opportunity to fill a gap in data available for research and greater geographic representation.

Specific Aim 2: A text field for occupation was available for a large proportion of patients (81.0%) and submitted for coding by the NIOCCS coding software. The software coded records, as compared to the manual coding, revealed high agreement using the kappa statistic even when including non-auto coded records (blanks). The amount of time taken for manual coding was shorter than expected. Due to several repeat records, coding time was approximately 30 minutes for 100 records of both industry and occupation. Auto-coding of 33,868 records took less than 12 hours from submission with no other users in the queue for coding. At the highest confidence, 37.4% of the records were left blank (not coded) by the software indicating an unrecognized text entry such as a business name or misspelling. Of the successfully coded records, 10,707 records (50.4%) were coded as “unknown” indicating that the patient’s usual occupation was recorded as “unknown” in the text of the record. Although only 62.6% of the records were successfully coded by NIOCCS at the best confidence level (90%), the coding software significantly reduced the amount of time to code several thousands of patient records. Thus, NIOCCS may drastically reduce the burden of manual coding industry and occupation in patient records and encourage information to be included in EHR systems in a more systematic way. These results indicate coding industry and occupation of patients from EHR records is feasible, generates results comparable to manual

coding, improves efficiency of assigning codes, and should add to the literature in accordance with the fourth IOM recommendation (13).

Specific Aim 3: Logistic regression analysis was conducted for hypertension diagnosis and elevated BP at the last clinic visit in 2012 for patients (Table 4). Factors largely associated with a hypertension diagnosis included sex, age, obesity, lipid disorder, and type 2 diabetes. Factors largely associated with elevated BP at the last patient visit included sex, age, and obesity. In both models, interaction was observed between age and obesity, and age and sex. Obese people and males were at a higher risk of being diagnosed with hypertension or having elevated BP, however, the effects of obesity and gender were reduced with aging. MSFW status was not significantly associated with either elevated BP or hypertension diagnosis. Much of the information gained from this regression supports known literature regarding hypertension (86), but little is previously published on whether specific vulnerable working populations, such as MSFWs, are at a higher risk of hypertension. Given that hypertension is such a widespread health concern and we found it to be a frequent diagnosis among this population, it is important to examine if this specific occupation is a risk factor for hypertension. Additionally, this study underlined the feasibility of EHR data to identify risk factors of hypertension from a medical perspective.

Specific Aim 4: Of the 61 laboratory values available from the EHRs, only 24 were matched with a baseline and exposed cholinesterase value. The 12 pairs showed depressed cholinesterase from baseline to suspected exposure, which could suggest some degree of pesticide exposure, especially when associated with compatible symptoms and signs.

Although limited, we were successful in extracting laboratory data from an EHR database to estimate exposure to certain pesticides. IOM recommendation 9 (Appendix A) suggests linking patient occupation and problem listings with an EHR may improve understanding of occupational exposure-disease relationships. In combination with meaningful use, including both occupation and laboratory indicators may improve the efficiency of diagnostic procedures and guide clinicians during diagnostic procedures (13). As a related example, if a patient presents with acute neurological symptoms, knowing that he is a farmworker should prompt the physician to test for abnormal cholinesterase and retest after exposure is reduced. Including laboratory results and other exposure indicators in the EHR can also address IOM recommendation 10, which included impacting meaningful use goals, such as incorporating laboratory data within the EHR, improving population health and reducing health disparities.

The results of this study are consistent with the sparse literature available on MSFWs (54,56,57,63). Most studies found similar demographics among MSFWs, especially the high percentage of Hispanic ethnicity and lower percentage of health insurance coverage. Nearly 25 years ago, type 2 diabetes and hypertension were among the most frequent diagnoses for MSFW patients at MHCs in Michigan, Indiana and Texas (56) but this study examined only MSFWs of all ages, rather than all adult patients by MSFW status. For hypertension, the 2005-2006 NHANES (88) estimated that 29% of U.S. adults were stage 1 hypertensive or higher (above 140/90). In our study, the prevalence for all patients was (20.1%) which was lower than the NHANES, but we did not have information about medication use or controlled hypertension. The information from this study may help prioritize preventive care

to improve the health outcomes in this working population. Finally, a major strength of this study was the opportunity to have a built in comparison group by having access to non-MSFW patients attending the same clinic.

Study Limitations

Missing data

A substantial number of data were missing for nearly half of all patients in 2012 and, depending on the distribution of missing values, could impact the interpretation of results. Due to the novelty of the data transfer process, it is unclear why there were substantial numbers of missing data. Missing values may be due to errors during the data transfer, to a lack of collection by clinic staff, or even refusal by the patient to provide information upon request. Handling EHR data has not been extensively studied in the literature. The initial data transfer was queried through a SQL server that organized the data by encounter, rather than by patient. Future databases generated from this HIE will be formatted by patient ID, rather than by encounter ID, or perhaps be in a flexible format that allows better manipulation for the researcher. Missing values for indicators may occur because the patient was not seen in 2012, or their visit may not have included the vitals.

Clinic staff may not be collecting the information because the EHR system does not prompt the provider for input. In a recent study of primary care clinics in Spain, researchers estimated only about one third of patient records had some indicators of occupation. The study examined both paper and electronic health records and found that after implementing an EHR system, occupation was only found in about 8% of patient records. The EHR system

did not have a specific field for occupation, demonstrating the first step in collecting occupation in an EHR is simply having the field available to the provider (122). A another study in a Spanish hospital system showed that recording of occupation differed between hospital departments, with Rehabilitation Services recording occupation more often than other hospital departments (123). In other cases, improvements in training of clinic staff may be the best way to improve the number of patients with informative occupation recorded.

Future analysis of complete clinic data may assist in understanding why many of the records contained missing data. A comparison between EHR data and data abstraction may serve to validate utilizing EHR as comparable to data abstraction and reduce concern of selection bias due to lack of data collection.

Misclassification Bias

In regard to the MSFW category, information about standard clinic practice was obtained from CBRN clinic and community partners in order to understand the function and reliability of this variable. The MSFW category is currently documented in the EMR system as a billing assignment. Instead, there should be a special population designation category, in addition to the billing field. This would allow the documentation of both the MSFW status and in the event that the patient has third party coverage, the billing information.

On a national basis, FQHCs designated as a MHC receive a portion of their annual grant as PHS 330 (g) funding in order to address the cost of care for MSFW patients. The amount of this grant is calculated on the basis of a projection of the number of patients to be seen and historically new start awards are made on an average of between $250 and $400 per

MSFW user per year, seldom enough to cover the cost of care for more than two or three medical encounters, let alone provision of dental, behavioral, or ancillary services such as outreach, transportation, interpretation, or environmental services, which are essential to serving this population. Therefore, if a MSFW qualifies for third party reimbursement, such as Medicare, the Colorado Indigent Care Program (CICP), or has commercial insurance, the clinic staff might not document a patient's MSFW status, thinking to shift the cost to that other payor in order to receive a higher reimbursement. This is probably not intentional, but merely a difference in eligibility.

Providers also track MSFW status using a non-payor related variable that was not transferred with this database but is critical in future analyses to ensure the distinction between the two identifiers and better understand the percentage of MSFWs that are utilizing the different payor systems (personal communications Ed Hendrikson, 4/18/2013, E. Roberta Ryder, 6/13/203). Using a non-payor related variable may reduce misclassification of MSFWs and improve understanding of MSFW medical needs.

Data Management and Quality

The quality of data received may also benefit from improved coding or consistency in collection. For example, information in the education variable was not systematically collected and was therefore difficult to analyze. The field was an open text field, rather than a prescribed selection. In the insurance field, there were specific options (self-pay, federal programs (Medicaid/Medicare), state programs (Colorado Indigenous Care Program, Children's State Insurance program), or private insurance. Obtaining patient level of

educational attainment can easily be conducted in a systematic coding scheme modeled after any national survey. Other examples of variables obtained but inconsistently coded included income, allergies, family histories, medication. Additionally, certain variables were collected but missing completely, such as smoking status. Smoking status is particularly important when examining hypertension (86). Including information on demographics is part of meaningful use standards, and refining these variables using coding schemes and improved quality will be important at later stages of meaningful use. Using information such as education, smoking status, family histories, etc., the provider can improve their care delivery and epidemiologists can improve recommendations for understanding health risk factors and future interventions. Beyond meaningful use, including demographics is important in epidemiologic research to understand health issues of MSFWs since few resources currently exist.

Generalizability

The records from this study come from a Colorado patient population and may not be generalizable to all MSFWs, FQHC users, general clinic patients or the general population. When conducting research in a working population, the healthy worker effect should also be considered as a potential selection bias where workers may be healthier than non-workers (124). In this case, MSFWs were more often reported as employed compared to non-MSFWs and may actually be healthier due to the healthy worker effect, although the impact may vary by health outcome. However, we cannot ascertain this from this study since this study gives a look into the feasibility of conducting health research in working groups using existing,

recent EHRs rather than conducting chart reviews. With improved industry and occupation information, comparisons of various working groups may reduce the potential selection bias from the healthy worker effect by comparing two working patient populations. Therefore, we just discuss the possibility of HWE reducing the burden of disease shown in the EHR.

Study Strengths

Generally speaking, health information technology is having a positive impact on care (125) but continued efforts are necessary to improve the usability and quality of EHR data. This study outlines some of the challenges of using EHR data and provides practical suggestions for improvement. This study also shows how quickly data can be transferred from the clinic to the HIE, de-identified, and transferred to the researcher (less than one year). However, the initial phase of assembling partnerships in the community, the clinic and academia can take several years to accomplish. As the CBRN was forming, much effort was taken to compile a medical database for research and found that the HIE provided the best foundation for the future of a national database.

The NIOCCS software is an excellent tool to decrease the burden of coding secondary industry and occupation data but cannot completely replace the insight of a trained manual coder. Manual coding remains necessary, specifically when a company name is listed requiring further investigation to understand the nature of the industry of that company and possibly the occupation. The current version of NIOCCS is the first available to the public and many enhancements are expected. Future versions of NIOCCS are expected to identify

and code information on job duties, specific businesses and reduce the amount of time needed for manual coding.

Including such laboratory values in a national HIE may improve follow-up in migratory populations. In order to achieve a second cholinesterase lab value, providers used outreach workers to locate and remind patients to return to the clinic. Follow-up with patients in obtaining laboratory values, such as cholinesterase, was difficult. However, certain labs should include information about the methodology since different methods of obtaining exposure may yield incomparable results. Other valuable information in order to address pesticide exposures include recording crops, job tasks, time at each crop/task and perceived exposures. While it may not be feasible for clinic staff to collect all of this information, future versions of the HIE may be linked to disease registries, poison control centers, and other databases to examine related trends.

Meaningful use seems to drive data quality. Stage 1, Step 5 meaningful use requirements include collection of demographics and vital signs during a patient visit. Core measure 5 involves collection of preferred language, gender, race, ethnicity, and date of birth. Core measure 8 involves collection of vital signs, including height, weight, blood pressure, and BMI calculation for adults. EHRs have the capacity to collect many different patient indicators, but the best quality (consistency in coding, and entry) was found in those required or directly related to meaningful use.

Future Research

The purpose of this study was to assess the feasibility of using EHR data as an epidemiologic research database of patients with the intention of future studies to include prospective, longitudinal research in MSFWs and other populations that utilize FQHCs. Such studies will include additional data collection, improved validation of data from the clinic to the researcher, and may require additional IRB approval and informed consent.

Future analysis of hypertension may include prevalent diagnoses of hypertension and understanding uncontrolled hypertension, where a diagnosis was made by the physician but elevated blood pressure persisted. Further analysis of hypertension should include information about smoking status, education, and disease covariates. As previously mentioned, smoking status was not included in the database which did not allow for control of this potential confounder in the regression model.

Recommendations for future monitoring of pesticide exposure include encouraging clinicians to monitor cholinesterase during the off season to establish a baseline where feasible since cholinesterase may be abnormally high during period of non-exposure because of repeat exposure (a rebound effect) and following required recommendations from surveillance programs. Cholinesterase depression is just one way to measure pesticide exposure and is specific to cholinesterase inhibiting chemicals such as carbamates and organophosphates (96). Perhaps it can be a model for other methods of including occupational exposures in the EHR.

Conclusion

The goals of this study were to examine the feasibility of using an EHR database to describe the patient population based on the specific aims including demographics, medical concerns, high blood pressure risk factors, and pesticide exposure; and to determine if coding of patient occupation and industry was feasible and informative from a text field within the EHR. From the FQHC database, we were able to successfully describe the overall patient population and compare known MSFWs to known non-MSFWs, but also had a significant issue with missing demographic and medical data. Our findings indicated that chronic disease is similar among this group of MSFW patients compared to other patients groups at this clinic. Also, our findings were consistent with the current literature on hypertension risk factors, including age, obesity and sex. Although limited, we were also able to extract laboratory cholinesterase values, which are a marker for certain pesticide exposure. A national effort is currently underway to improve health outcomes by switching to EHR systems, a much needed enhancement. This may be an efficient way to improve epidemiologic research in populations with a weak knowledge base, specifically MSFWs. Several agencies are currently charged with improving health research by tapping into the new resources of available data in EHRs through HIEs. This innovate source of data may be the next major movement to improve injury and illness in relation to occupation.

FIGURES & TABLES

Figure 1: Electronic patient record inclusion diagram (Salud Family Health Center, 2012)

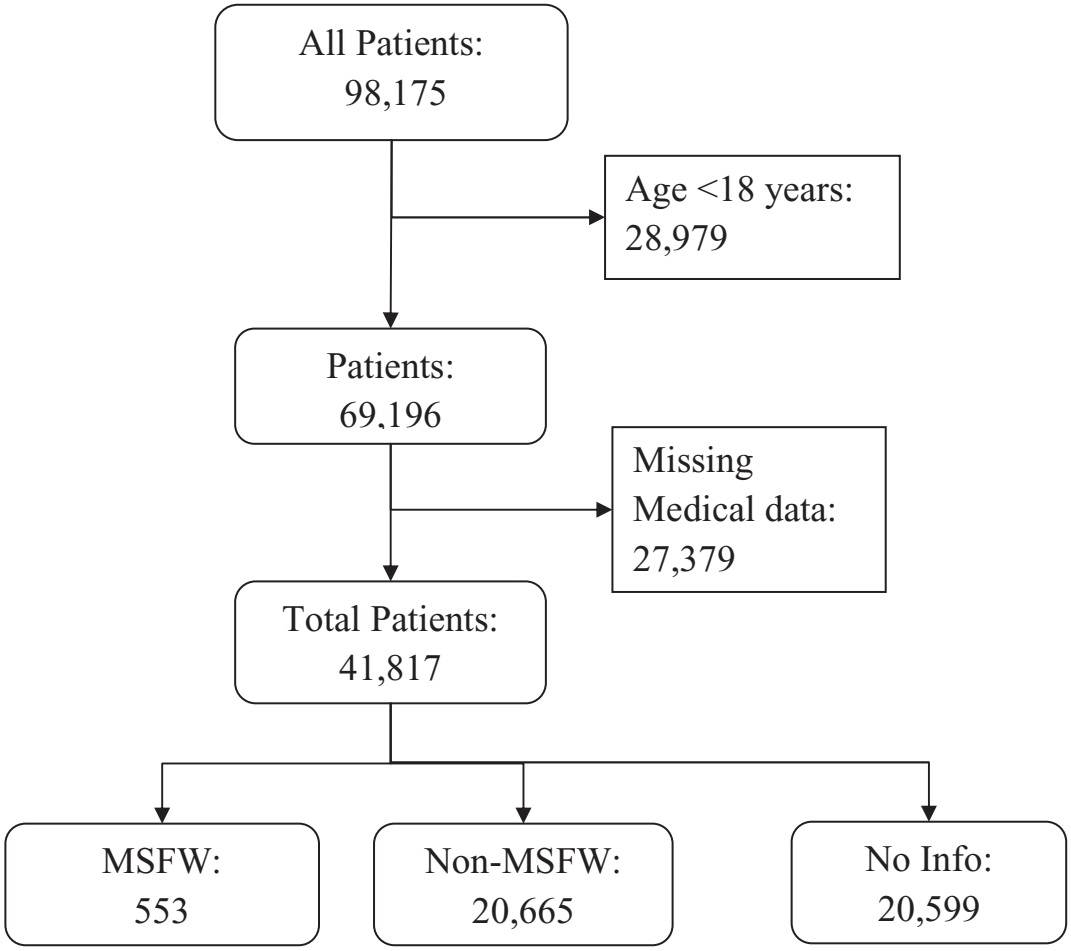

Table 1: Demographics of all patients and patients by MSFW status (Salud Family Health Center, 2012)

	All Patients	MSFW	Non-MSFW	No Info	P-Value[a]
Number of patients	**41,817**	**553**	**20,665**	**20,599**	
Age, years					
median (IQR)	39 (28,52)	39 (29,51)	44 (31,55)*	36 (26,48)*	<0.001
Total patients (N)	41,817	553	20,665	20,599	
Sex					
Males, n (%)	15,930 (38.5)	280 (50.6)	7,713 (37.3)*	7,937 (39.3)*	<0.001
Total patients (N)	41,421	553	20,665	20,203	
Marital Status					
Married, n (%)	8,598 (41.2)	306 (56.6)	8,242 (40.8)*	50 (42.0)*	<0.001
Total patients (N)	20,847	541	20,187	119	
Employment					
Employed, n (%)	6,659 (34.5)	174 (46.4)	6,451 (34.2)*	34 (39.1)	<0.001
Total patients (N)	19,302	375	18,840	87	
Language					
English, n (%)	14,231 (67.0)	124 (22.5)	14,063(68.3)*	44 (57.1)*	<0.001
Total patients (N)	21,233	552	20,604	77	

	All Patients	MSFW	Non-MSFW	No Info	P-Value[a]
Public housing					
Yes, n (%)	422 (2.0)	6 (1.1)	414 (2.0)	2 (2.7)	0.23
Total Patients (N)	21,131	539	20,519	73	
Homeless					
Yes, n (%)	691 (3.3)	6 (1.1)	685 (3.3)*	-	0.002
Total patients (N)	21,214	549	20,665		
Race/Ethnicity, n (%)					
All Hispanic	21,172 (53.7)	452 (85.8)	9,704 (49.1)*	11,016 (57.6)*	
Non-Hisp.-White	16,020 (40.6)	70 (13.3)	8,984 (45.4)	6,966 (36.4)	<0.001
Non-Hisp.- other	2,260 (5.73)	5 (0.95)	1,095 (5.5)	1,160 (6.1)	
Total patients (N)	39,452	527	19,783	19,142	
Insurance, n (%)					
Self-Pay	734 (4.0)	175 (42.6)	556 (3.1)*	3 (3.0)*	
Federal/State	15,027 (81.7)	209 (50.9)	14,746 (82.5)	72 (72.7)	<0.001
Private Insurance	2,626 (14.3)	27 (6.6)	2,575 (14.4)	24 (24.2)	
Total Patients (N)	18,387	411	17,877	99	

a=Overall difference using χ^2 test (categorical variables) or Kruskal-Wallis test (continuous variable, age)
*=Significantly different compared to the MSFW group using Dunn Procedure (Kruskal-Wallis), Fisher's Exact test, or $\chi 2$ test with the Bonferroni correction for multiple comparisons.

Table 2a: Description of medical indicators (Salud Family Health Center, 2012)

	All Patients	MSFW	Non-MSFW	No Info	P-value[a]
Number of patients	**41,817**	**553**	**20,665**	**20,599**	
Obesity					
BMI>=30, n (%)	14,832 (39.3)	227 (42.5)	8,525 (43.9)	6,080 (34.3)*	<0.001
Total Patients (N)	37,702	534	19,433	17,735	
BMI, kg/m2					
mean (SD)	29.3 (7.0)	29.7 (5.8)	30.1 (7.3)	28.5 (6.6)*	<0.001
Total Patients (N)	37,702	534	19,433	17,735	
Systolic BP, mmHg					
median (IQR)	120 (110, 130)	122 (112,132)	120 (110,128)*	120 (110,132)*	<0.001
Total Patients (N)	41,554	551	20,605	20,398	
Diastolic BP, mmHg					
median (IQR)	78(70,82)	78(70,84)	79 (70,84)*	77(70,82)*	<0.001
Total Patients (N)	41,554	551	20,605	20,398	
BP Categories, n (%)					
Normal	15,213 (36.6)	180 (32.7)	6,667 (32.4)	8,366 (41.0)*	<0.001
Pre-hypertension	18,017 (43.4)	257 (46.6)	9,058 (44.0)	8,702 (42.7)	
Hypertension Stage 1	6,220 (15.0)	81 (14.7)	3,519 (17.1)	2,620 (12.8)	
Hypertension stage 2	2,105 (5.1)	33 (6.0)	1,361 (6.6)	710 (3.5)	
Total Patients (N)	41,554	551	20,605	20,398	

	All Patients	MSFW	Non-MSFW	No Info	P-value[a]
Heart Rate, beats/min					
median (IQR)	80 (72,84)	77 (70, 84)	80 (72, 86)*	78 (72, 84)*	<0.001
Total Patients (N)	40,143	549	20,333	19,261	
Temperature, ◦F					
mean (SD)	98.03 (0.6)	98.07 (0.6)	98.01 (0.6)	98.06 (0.6)	<0.001
Total Patients (N)	37,381	526	18,683	18,172	
Past Surgery					
Yes, n (%)	27,657 (66.1)	311 (56.2)	15,126 (73.2)*	12,220 (59.3)*	<0.001
Total Patients (N)	41,817	553	20,665	20,599	

*=(Pairwise)Significant difference compared to MSFWs based on Fisher's Exact test, χ^2 test, ANOVA or Dunn Procedure (Kruskal-Wallis)
a=Overall difference using χ^2 test (categorical) or Kruskal-Wallis test (continuous)

Table 2b: Top 10 most common diagnoses among patients with available diagnosis codes, excluding diagnostic or preventive codes (Salud Family Health Center, 2012)

	All Patients	MSFW	Non-MSFW	No Info	P-value[a]
Number of patients	**41,817**	**553**	**20,665**	**20,599**	
Prevalence of Top Diagnoses[b], n (%)					
Hypertension	8,087 (19.3)	122 (22.1)	5,860 (28.4)*	2105 (10.2)*	<0.001
Overweight/Obese	6,047 (14.5)	190 (34.4)	4,543 (22.0)*	1,314 (6.4)*	<0.001
Lipid Disorder	5,604 (13.4)	77 (13.9)	4,069 (19.7)*	1,458 (7.1)*	<0.001
Type 2 Diabetes	4,954 (11.9)	73 (13.2)	3,416 (15.5)	1,465 (7.1)*	<0.001
Back disorder	4,861 (11.6)	68 (12.3)	1,576 (7.5)	3,217 (15.6)*	<0.001
Anxiety	4,434 (10.6)	44 (8.0)	3,203 (15.5)*	1,187 (5.8)	<0.001
Joint Disorder	3,735 (8.9)	44 (8.0)	2,482 (12.0)*	1,209 (5.9)	<0.001
Abdominal Pain	3,591 (8.6)	41 (7.41)	2,145 (10.4)*	1,405 (6.8)	<0.001
Drug Abuse	3,519 (8.42)	47 (8.5)	2,741 (13.3)*	731 (3.6)*	<0.001
Soft Tissue Disorder	3,158 (7.55)	39 (7.05)	2,080 (10.1)*	1,039 (5.0)	<0.001

*-Significantly difference compared to MSFWs based on Fisher's Exact test
a-Overall difference using χ^2 test (categorical)
b-Each patient may have had more than one diagnoses, but multiple diagnoses of the same health outcome were removed. The top 10 diagnoses were based on all patients rather than each individual group

Table 3: NIOCCS Auto coded records at various confidence levels and agreement of manually coded records compared to auto-coded records (Salud Family Health Center, 2012)

	Number of Records	Percent coded (Industry & Occupation)	Kappa Industry	Kappa Occupation	Kappa Industry (without blanks)	Kappa Occupation (without blanks)
	N=33,868		n=100	n=100		
Successful coded records at:						
30% confidence	26,742	79.0%	0.639	0.646	0.803 n=83	0.804 n=83
70% confidence	24,435	72.1%	0.639	0.634	0.910 n=75	0.897 n=75
90% confidence	21,212	62.6%	0.580	0.575	1.000 n=65	0.979 n=65

Table 4: Risk factors[a] significantly associated with hypertension or elevated blood pressure (>120/80mmHg), (Salud Family Health Center, 2012)

	Hypertension Diagnosis Odds Ratio (95% CI)	Elevated BP at last Visit Odds Ratio (95% CI)
	n=16,524	n=15,578
MSFW[b]	0.80 (0.59, 1.09)	1.09 (0.81, 1.47)
Sex (male)	5.37 (3.89, 7.43)	8.23 (6.40, 10.58)
Age	1.09 (1.08, 1.09)	1.05 (1.04, 1.05)
Race/Ethnicity		
Hispanic	1	
Non-Hispanic White	0.98 (0.88, 1.09)	-
Other	1.17 (0.97, 1.40)	
Insurance[b]		
Self Pay	1	1
State/Federal Program	0.89 (0.69, 1.14)	1.21 (0.99, 1.49)
Private Insurance	0.96 (0.73, 1.25)	1.14 (0.91, 1.42)
English	1.12 (0.99, 1.26)	1.23 (1.12, 1.35)
Employed	-	1.08 (1.00, 1.17)
Surgery	1.23 (1.11, 1.36)	-
Obesity	3.61 (2.61, 5.00)	4.35 (3.42, 5.54)
Lipid Disorder	2.07 (1.89, 2.27)	1.14 (1.03, 1.28)
Type 2 Diabetes	2.43 (2.20, 2.68)	1.22 (1.08, 1.37)
Anxiety	1.32 (1.18, 1.47)	-
Drug Abuse	1.16 (1.04, 1.30)	-
Soft Tissue Disorder	1.18 (1.04, 1.33)	-
Obesity*Age	0.99 (0.98, 0.996)	0.99 (0.98, 0.995)
Sex* Age	0.98 (0.97, 0.98)	0.97 (0.97, 0.98)

a. We compared MSFW to Non-MSFW in this analysis (excluded the no-info group). Homeless status, and Public Housing Status were excluded from the model due to missing values: The full model included language, marital status, employment status, presence of surgical history, diagnoses for the presence of back disorder, presence of joint disorder, presence of abdominal pain, and presence of drug abuse. Variables included above were included due statistical significance (P<0.10) or clinical relevance

b. Included in the final-model but no significant association (P> 0.10)

Table 5: Cholinesterase values for 12 MSFW patients (Salud Family Health Center, 2012)

Patient Number	Baseline Cholinesterase (IU/L)	Exposed Cholinesterase (IU/L)	% Difference
1	3726	3657	-1.9%
2	5785	5117	-11.5%
3	5122	4398	-14.1%
4	4645	4363	-6.1%
5	4923	4616	-6.2%
6	5776	4970	-14.0%
7	5961	5561	-6.7%
8	5216	5029	-3.6%
9	5068	4637	-8.5%
10	4751	4643	-2.3%
11	5028	4699	-6.5%
12 (male)	3119*	3117*	-0.1%

*Below range of normal (Normal= 3342-7586 IU/L for males; 2673-6592 IU/L for females)

APPENDICES

Appendix A: Institute of Medicine Recommendations Regarding incorporating Occupation and Industry (13)

"Initial Focus on Occupation, Industry, and Work-Relatedness Data Elements

1. Conduct Demonstration Projects to Assess the Collection and Incorporation of Information on Occupation, Industry, and Work-Relatedness in the EHR
   NIOSH, in conjunction with other relevant organizations and initiatives, such as the Public Health Data Standards Consortium and Integrating the Healthcare Enterprise (IHE) International, should conduct demonstration projects involving EHR vendors and health care provider organizations (diverse in the services they provide, populations they serve, and geographic locations) to assess the collection and incorporation of occupation, industry, and work-relatedness data in the EHR at different points in the workflow (including at registration, with the medical assistant, and with the clinician). Further, to examine the bidirectional exchange of occupational data between administrative databases and clinical components in the EHR, NIOSH in conjunction with IHE should conduct an interoperability-testing event (e.g., Connectathon) to demonstrate this bidirectional exchange of occupational information to establish proof of concept and, as appropriate, examine challenges related to variable sources of data and reconciliation of conflicting data.

2. Define the Requirements and Develop Information Models for Storing and Communicating Occupational Information
   NIOSH, in conjunction with appropriate domain and informatics experts, should develop new or enhance existing information models for storing occupational information, beginning with occupation, industry, and work-relatedness data and later focusing on employer and exposure data. The information models should consider the various use cases in which the information could be used and use the recommended coding standards. consider how best to use social history templates to collect a work history and the problem list to document exposures and abnormal findings and diagnoses with optional work-associated attributes for possible, probable, or definite causes; exposures; and impact on work.

3. Adopt Standard Occupational Classification (SOC) and North American Industry Classification System (NAICS) Coding Standards for Use in the EHR
   NIOSH, with assistance from other federal agencies, organizations, and stakeholders (e.g., Bureau of Labor Statistics, Census Bureau, Council of State and Territorial Epidemiologists [CSTE], National Library of Medicine, National Institute of

Standards and Technology, National Uniform Billing Committee, Health Level 7 International [HL7]), should recommend to the Health Information Technology (IT) Standards Committee the adoption of SOC and NAICS to code occupation and industry. Furthermore, NIOSH should develop models for reporting health data from EHRs by occupation and industry at different levels of granularity that are meaningful for clinical and public health use.

4. Assess Feasibility of Autocoding Occupational Information Collected in Clinical Settings
   NIOSH should place high priority on completing the feasibility assessment of auto-coding the narrative information on occupation and, where available, industry that currently is collected and recorded in certain clinical settings, such as the Dartmouth-Hitchcock health care system, Kaiser Permanente, New York State Occupational Health Clinic Network, Cambridge Health Alliance, and hospitals participating in the National Electronic Injury Surveillance System.

5. Develop Meaningful Use Metrics and Performance Measures
   Based on findings from the various demonstration projects and feasibility studies, NIOSH, with the assistance of relevant professional organizations and the Health IT Policy Committee, should develop meaningful use metrics and health care performance measures for including occupational information in the meaningful use criteria, beginning with the incorporation of occupation, industry, and work-relatedness data, and later expanding as deemed appropriate to include other data elements such as exposures and employer.

6. Convene a Workshop to Assess Ethical and Privacy Concerns and Challenges Associated with Including Occupational Information in the EHR
   NIOSH should convene a workshop involving representatives of labor unions, insurance organizations, health care professional organizations, workers' compensation-related organizations (e.g., International Association of Industrial Accident Boards and Commissions, National Council on Compensation Insurance), and EHR vendors to
   - assess the implications for the patient and clinician of incorporating work-relatedness in the EHR, with respect to workers' compensation; and
   - propose guidelines and policies for protecting the patient's non-work-related health information from inadvertent disclosure and to ensure compliance with the Health Insurance Portability and Accountability Act, workers' compensation, and other privacy standards.

Enhance the Value and Use of Occupational Information in the EHR

7. Develop and Test Innovative Methods for the Collection of Occupational Information for Linking to the EHR
   NIOSH should initiate efforts in collaboration with large health care provider organizations, health insurance organizations, EHR vendors, and other stakeholders to develop and test methods for collecting occupational data from innovative sources. Specifically, NIOSH should evaluate collection methods that involve
   - patient input through mechanisms such as web-based portals and personal health records; and
   - other means such as health-related smart cards, health insurance cards, and human resource systems.

8. Develop Clinical Decision-Support Logic, Education Materials and Return-to-Work Tools
   NIOSH, relevant professional organizations, and EHR vendors should begin to develop, test, and iteratively refine and expand clinical decision-support tools for common occupational conditions(e.g., work-related asthma);
   - tools and programs that could be easily accessed for education of patients and caregivers about occupational illnesses, injuries, and workplace safety;
   - training modules for administrative staff to collect occupational information in different care settings; and
   - tools to improve and standardize functional job assessment and return-to-work documentation in EHRs, including standards for the transmission of these forms.

9. Develop and Assess Methods for Collecting Standardized Exposure Data
   NIOSH should continue to work with occupational and environmental health clinics and other relevant stakeholders to develop and assess methods for collecting standardized exposure data for work-related health conditions. NIOSH should explore the feasibility of listing possible or probable exposures in the problem list or elsewhere in the EHR;
   - linking occupational information in the EHR to online occupational, toxicological, and hazardous materials databases, such as the Occupational Information Network (O*NET), the Association of Occupational and Environmental Clinics, and Haz-Map, to enhance diagnosis and treatment of work-related illnesses and injuries; and
   - automatically generating codes for exposures based on narrative text entries.

10. Assess the Impact of Incorporating Occupational Information in the EHR on Meaningful Use Goals

    NIOSH, in conjunction with relevant stakeholders (e.g., Public Health Data Standards Consortium, CSTE, Association of State and Territorial Health Officials), should

    - develop measures and conduct periodic studies to assess the impact of integrating occupational information in EHRs, and
    - estimate the economic impact of EHR-facilitated return-to-work practices for both work-related and non-work-related conditions."

Appendix B: Protected Health Information (The Health Insurance Portability and Accountability Act of 1996 (HIPAA) Privacy and Security Rules)

1. Names
2. All geographic subdivisions smaller than a state, including street address, city, county, precinct, zip code and equivalent geocodes.
3. All elements of dates (except year) for dates directly related to an individual, including birth date, admission date, discharge date, date of death
4. Telephone numbers
5. Fax numbers
6. E-mail addresses
7. Social security numbers (although using the last 4 digits is acceptable)
8. Medical record numbers
9. Health plan beneficiary numbers
10. Account numbers
11. Certificate/license numbers
12. Vehicle identifiers, including license plate and serial numbers
13. Device identifiers and serial numbers
14. Web universal resource locators (URLs)
15. Internal protocol (IP) address numbers
16. Biometric identifiers, including finger and voice prints
17. Full-face photographic images and any comparable images
18. Any other unique identifying number, characteristic, or code

Appendix C: Data Use Agreement with the Community-Based Research Network

Data Use Agreement (DUA) Community Based Research Network (CBRN)

Purpose of Agreement: To provide for the use of data under specific conditions as described herein. Parties to this agreement, hereby enter into this Data Use Agreement ("DUA") as of February 1, 2013.

Parties to this Agreement:
Data Provider: E. Roberta Ryder, representing the Community Based Research Network (CBRN), (1770, FM 967, Buda TX 78610, 512- 312-5453, Ryder@ncfh.org) is the Data Provider.

Recipient Principal Investigator (PI): Christina Socias, (The University of Texas School of Public Health, 7411 John Smith Blvd, Suite 1100, San Antonio, TX 78229, 305-794-3031, christina.m.socias@uth.tmc.edu or socias1@gmail.com) is the Recipient Principal Investigator (PI).

Provider Summary: CBRN (funded by National Institute of Environmental Health Sciences (NIEHS) Grant # 1RC4ES019405) operates the National Farmworker Clinical Research Database ("the database"), as a Health Information Exchange (HIE) of patient Electronic Health Records (EHRs) with a goal to be a bidirectional and egalitarian community-academic partnership. Multiple Principal Investigators (PIs) are necessary to achieve such a partnership. The primary objective of the database is to collect existing clinical information related to Migrant and Seasonal Farmworkers (MSFWs) and their peers. The database will be managed and maintained by Centex Systems Support Services (CSSS) as part of the Integrated Care Collaborative, under the guidance of multiple PIs, including E. Roberta Ryder (CEO of the National Center for Farmworker health), Sharon Cooper (Regional Dean, UT Houston SPH, San Antonio Campus), and Edward Hendrikson (PA, & PhD Researcher, Salud Family Health Centers). All health centers contributing data to the HIE have established a Business Associate Agreement with CSSS in order to securely share patient EHR records. Participating Health Centers are designated Federally Qualified Health Centers (FQHC) serving Agricultural Workers and supported through funding from the United States Department of Health and Human Services (DHHS), Health Resources and Services Administration (HRSA). The goal of the CBRN is to develop the infrastructure and establish a national HIE that will collate current clinical data on MSFWs and their peers from selected participating health centers. To accomplish CBRN's objectives and goals, an HIE database of EHRs from participating Health Centers has been constructed to share with qualified researchers for secondary data analysis after receiving approval by CBRN's Steering and Advisory Committees of the terms of this DUA.

The Database: The database is assembled as a HIE and will be shared as a "limited database" as defined by the Health Insurance Portability and Accountability Act (HIPAA). It has been stripped of most personal identifiers. The health information that may remain in the database if requested by the Recipient PI includes:

1. dates such as admission, discharge, service, DOB, DOD;
2. city, state, five digit or more zip code; and
3. ages in years, months or days or hours.

Data Use Agreement (DUA) Community Based Research Network (CBRN)

The wealth of data available in individual records might make possible the individual identification of some subjects. To protect the confidentiality and privacy of these study participants, the Recipient PI, who is granted access to data, must adhere to the requirements of this DUA. Failure to comply with this DUA could result in denial of further access to the database and may leave requesting investigators liable to legal action on the part of study participants, their families, or the U. S. Government. Recipient PI, Christina Socias, requests access to the database at the Recipient PI's sole risk and at no expense to CBRN and/or NIEHS.

Summary of DUA Requirements as detailed in the "Terms and Conditions." The recipient PI will:

1. Forward IRB approval to CBRN to initiate sharing of data
2. Notify CBRN of any substantive changes in the Research Project
3. Notify CBRN within 60 days of publication; and
4. Notify CBRN when data has been destroyed.

Data Use Agreement (DUA) Community Based Research Network (CBRN)

AGREED TERMS AND CONDITIONS

It is mutually agreed as follows:
The database will be used by Recipient PI solely in connection with the Research Project outlined below, specifically described in an attached Appendix A. Data will be accessed from Institutional servers and devices only. At no time should any data be transferred or copied to a personal server, computer or device.

The Research Project description includes: project title, a 1-2 paragraph summary description of the study aims, design, and a brief description of the analysis plan. Recipient PI agrees that the database will not be used in any research that is not disclosed and approved as part of the Research Project.

The Recipient PI also agrees to provide the CBRN Steering/Advisory Committee with proof of Institutional Review Board ("IRB") approval, and to attempt to incorporate suggested changes into their Research Project. Recipient PI agrees to report promptly to CBRN any proposed change in the Research Project and any unanticipated problems involving risks to subjects or others. This agreement is made in addition to, and does not supersede any of Recipient PI's institutional policies or any local, State, and/or Federal laws and regulations, which provide additional protections for human subjects.

For purposes of this Research Project ("The Meaningful use of Incorporating Occupation and Industry in an Electronic Health Record (EHR) Database for Epidemiologic Research"), The Recipient PI agrees to the following:

1. **Research Project**. This DUA covers only the Research Project described in Exhibit A "The Meaningful use of Incorporating Occupation and Industry in an Electronic Health Record (EHR) Database for Epidemiologic Research". Recipient PI will submit a separate completed DUA for each research project for which data may be requested in the future.

2. **Non-transferability**. This DUA is not transferable. Recipient PI agrees that substantive changes made to the Research Project described above, and/or appointment by Recipient or another Recipient Principal Investigator to complete the Research Project, will require execution of a new DUA in which the new Recipient Principal Investigator and/or new Research Project are designated. Furthermore, Recipient agrees to retain control over database, and further agrees not to transfer Repository Data to any other entity or any individual. All data should remain on recipient institution servers and/or authorized IT equipment, computers or other devices. Database should not be transferred in whole or part to any personal computer devices.

3. **Publication**. Recipient PI agrees to provide to CBRN a copy of any peer-reviewed published article, meeting abstracts and presentations for project records within 60 days of publication.

4. **Acknowledgments**. Recipient PI agrees to acknowledge the contribution of the CBRN in any and all oral and written presentations, disclosures, and publications resulting from any and all analyses of Repository Data. The names of CBRN investigators and their institutional roles as described in the attached letters of support (Appendix B). Acknowledgement of the contribution of CBRN is expected in all oral and written presentations. Recipient PI will acknowledge the source of the database by including language similar to the following either in the acknowledgments or in the text of the manuscript:

   > "This manuscript was prepared using a limited access dataset obtained from the Community Based Research Network; however manuscript does not necessarily reflect the opinions or views of the Community Based Research Network or National Institute of Environmental Health Sciences".

5. **Collaborations**. If the Research Project involves collaboration with CBRN Investigators or committee members, then the manuscript will also be reviewed by those contributing authors and Recipient PI will provide co-authorship to contributing Investigators and/or members, if scientifically appropriate.

6. **Non-Identification**. Recipient PI agrees that data will not be used, either alone or in conjunction with any other information, in any effort whatsoever to establish the individual identities of any of the subjects from whom data were obtained.

7. **Non-Data**. Notwithstanding the definition of "database" or the agreed Terms and Conditions of this DUA, Recipient PI's obligations under this DUA shall not extend to any information:
   (a) that can be demonstrated to have been publicly known at the time of disclosure; or
   (b) that can be demonstrated to have been in the possession of or that can be demonstrated to have been readily available to Recipient PI from another source prior to the disclosure; or
   (c) that becomes part of the public domain or publicly known by publication or otherwise, not due to any unauthorized act by Recipient PI; or
   (d) that can be demonstrated as independently developed or acquired by Recipient PI without reference to or reliance upon database provided under this Agreement; or
   (e) that is required to be disclosed by law, provided the Recipient PI takes responsible and lawful actions to avoid and/or minimize such disclosure.

8. **Non-Endorsement, Indemnification**. Recipient PI agrees not to claim, infer, or imply Governmental endorsement of the Research Project, the entity, or personnel conducting the Research Project or a resulting commercial product(s) except as described in section 4. It is the intention of the Recipient PI that the Data Provider not be liable to any parties

for any liabilities, demands, damages, expenses, or losses arising from the Recipient PI's use for any purpose of Data. No indemnification is provided or intended by either Party.

9. **Amendments.** Amendments to this DUA must be made in writing and signed by authorized representatives of all parties.

10. **Termination.** CBRN may terminate this DUA if Recipient is in default of any condition of this DUA and such default has not been remedied within 30 days after the date of written notice by CBRN's authorized representative of such default.

11. **Disqualification, Enforcement.** Failure to comply with any of the terms specified herein may result in disqualification of Recipient PI from receiving additional Repository Data. The United States Government shall have the right to institute and prosecute any proceeding at law or in equity against the Recipient PI for violating or threatening to violate the confidentiality requirements of this agreement, the limitations on the use of the Data provided, or both. Proceedings may be initiated against the violating party, legal representatives, and assigns, for a restraining injunction, compensatory and punitive damages, mandamus, and/or any other proceeding in law or equity, including obtaining the proceeds from any intellectual property or other rights that are derived in whole or in part from the breach of the confidentiality requirements or use limitations of this agreement. In addition, Recipient PI acknowledges and agrees that a breach or threatened breach of the confidentiality requirements or use limitations of this agreement may subject Recipient PI to legal action on the part of database subjects, their families, or both.

12. **Project Completion.** After publication of all materials (articles, abstracts, presentations, etc.) the Recipient PI agrees to remove/destroy all data acquired from the Data Provider. No data in whole or in part should remain on any institutional servers or devices. Further, Recipient PI agrees to notify CBRN when that data has been destroyed.

Data Use Agreement (DUA) Community Based Research Network (CBRN)

The undersigned Authorized Representative of the Data Provider and Recipient PI expressly certify and affirm that the contents of any statements made herein are truthful and accurate. The authorized representative of CBRN has been appointed unanimously by all CBRN committee members and whose signature below represents the approval of the Data Provider.

RECIPIENT PI INFORMATION AND SIGNATURE

Signature of Recipient PI — Date: 2/4/13

Name of Recipient PI: Christina Socias, DrPH (candidate)
Title of Recipient PI: Graduate Research Assistant
Email of Authorized Official: Christina.M.Socias@uth.tmc.edu
Recipient Institution and address: University of Texas School of Public Health San Antonio Regional Campus
7411 John Smith Drive, Suite 1100
San Antonio, TX 78229

****

DATA PROVIDER INFORMATION AND AUTHORIZED REPRESENTATIVE'S SIGNATURE

Signature of Data Provider's Authorized Representative — Date: 02.04.13

Name of Authorized CBRN Representative: E. Roberta Ryder
Title of Authorized CBRN Representative: CEO, National Center for Farmworker Health
Email of Authorized Representative: Ryder@ncfh.org
Recipient Institution and address: Community Based Research Network
C/O National Center for Farmworker Health
1770, FM 967
Buda TX 78610

PLEASE RETURN SIGNED FORM AS A PDF TO:
E. Roberta Ryder at ryder@ncfh.org
and cc: Lisa Mendoza Miller at miller@ncfh.org

Data Use Agreement (DUA) Community Based Research Network (CBRN)

Appendix A: Summary of Research Project

Research Project title: "The Meaningful use of Incorporating Occupation and Industry in an Electronic Health Record (EHR) Database for Epidemiologic Research"

Specific Aims:

This project will-

1. Describe the patients that visited a MHC including demographic variables, social history, medical history, and medical diagnoses.
2. Assess the feasibility of coding occupation/industry text fields using the NIOSH Industry and Occupation Computerized Coding System (NIOCCS) and implement coding of industry and occupation text fields.
3. Describe the prevalence of hypertension diagnoses in MSFW patients compared to non-MSFW patients.
4. Describe pesticide exposure and related symptoms from a nested feasibility study that involved collection of serum cholinesterase lab results on approximately 100 MSFW patients.

Design and Analysis plan:

Few extensive clinical databases exist on the health of migrant and seasonal farmworkers (MSFWs). Electronic health records (EHRs) are increasingly utilized by Federally Qualified Health Centers (FQHCs) and have the potential to improve clinical care, as well as supplement current surveillance and epidemiologic studies of underserved populations, such as MSFWs. The purpose of this study is to examine the feasibility of using an EHR database in a descriptive epidemiologic study of a vulnerable working population.

Patients will be described based on demographics, social history, medical history, and diagnoses from a de-identified EHR database from a clinic that serves as a Migrant Health Center (MHC). Where available, patient industry and occupation will be coded using the U.S. Census coding scheme to assess the feasibility of retroactively coding occupation in text form within EHR databases. An analysis will be conducted to examine the prevalence of hypertension (ICD-9-CM = 401.X-405.X) by MSFW status. Approximately 100 MSFW patients were tested for pesticide exposure by serum cholinesterase before and after farm work seasons. An analysis will be conducted in the pilot subset to describe any change in cholinesterase from baseline (before farm work) and any reported symptoms controlling for other variables. We expect that the results of this study will add to the published literature on MSFWs and generate hypotheses about health risk factors associated with this population.

Data Use Agreement (DUA) Community Based Research Network (CBRN)

Appendix B: Letters of Support attached from CBRN Principal Investigators

- Sharon P. Cooper, PhD
  Academic and Administrative Contact Principal Investigator
  Professor and Regional Dean
  University of Texas Health Science Center School of Public Health

- E. Roberta Ryder, CEO
  Community Contact Principal Investigator
  National Center for Farmworker Health

- Edward Hendrikson, Ph.D., PA-C.
  Medical Clinic Contact Principal Investigator
  Environmental Health Director and physician assistant, Salud Family Health Center, Colorado.

Appendix D- Contingency tables for agreement between manual and software coded industry and occupation records at each confidence level.

Manually-coded industry compared to auto-coded industry (30% Confidence)

Manual Industry Codes	NIOCCS Generated Industry Codes																											
	018	029	077	118	317	349	357	396	547	559	579	699	739	769	789	827	856	866	868	877	919	947	988	989	999	blank	Total
017		1																									1
077			4																								4
118				1																							1
119																										1	1
336					1																						1
339							1																				1
349																										1	1
498																	1										1
547									1																		1
579											1																1
629																										1	1
639																										1	1
657						1																					1
669													1														1
769										1				2													3
777																										2	2
798								1																			1
817												1															1
819																					1					1	2
827																1											1
868	1																		6							1	8
877																				1							1
929																										1	1
947																						1					1
979																		1									1
988																							3				3
989																								17			17
999															1										32	8	41
Total	1	1	4	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	6	1	1	1	3	17	32	17	100

Manually-coded industry compared to auto-coded industry (70% Confidence)

Manual Industry Codes		NIOCCS Generated Industry Codes																		
		029	077	357	389	547	579	639	739	769	789	827	868	877	947	988	989	999	blank	Total
	017	1																		1
	077		4																	4
	118																		1	1
	119																		1	1
	336																		1	1
	339			1																1
	349																		1	1
	498																		1	1
	547					1														1
	579						1													1
	629																		1	1
	639							1												1
	657																		1	1
	669								1											1
	769									2									1	3
	777																		2	2
	798																		1	1
	817																		1	1
	819																		2	2
	827											1								1
	868												6						2	8
	877													1						1
	929																		1	1
	947														1					1
	979																		1	1
	988															3				3
	989																17			17
	999				1						1							32	7	41
	Total	1	4	1	1	1	1	1	1	2	1	1	6	1	1	3	17	32	25	100

Manually-coded industry compared to auto-coded industry (90% confidence)

Manual Industry Codes		NIOCCS Generated Industry Codes											
		077	547	579	769	827	868	877	988	989	999	blank	total
	017											1	1
	077	3										1	4
	118											1	1
	119											1	1
	336											1	1
	339											1	1
	349											1	1
	498											1	1
	547		1										1
	579			1									1
	629											1	1
	639											1	1
	657											1	1
	669											1	1
	769				2							1	3
	777											2	2
	798											1	1
	817											1	1
	819											2	2
	827					1							1
	868						4					4	8
	877							1					1
	929											1	1
	947											1	1
	979											1	1
	988								3				3
	989									17			17
	999										32	9	41
	Total	3	1	1	2	1	4	1	3	17	32	35	100

Comparison of Manual Occupation Codes to Auto-Coded Occupation (30% Confidence)

Manual Occupation Codes	NIOCCS Generated Occupation Codes																																
		016	031	062	071	263	284	313	365	394	411	413	422	423	476	605	620	626	701	720	772	825	876	880	901	905	906	910	913	990	blank	Total
	031		1																													1
	263					1																										1
	360						1																								1	2
	365								1																							1
	394									1																						1
	402																														2	2
	411										3																				1	4
	414											1																				1
	422												1																			1
	423							1						2																		3
	425																														2	2
	461																														1	1
	476														1																	1
	540																														1	1
	570		1																												1	2
	586																						1									1
	605															1																1
	626																	3														3
	720																			1												1
	739				1																											1
	772																				1											1
	775																														1	1
	780																														1	1
	781																							1								1
	901																								11							11
	905																									4						4
	906			1																							3					4
	910																											2				2
	913																												1			1
	962																														1	1
	990	1	1														1		1			1								32	5	42
	Total	1	3	1	1	1	1	1	1	1	3	1	1	2	1	1	1	3	1	1	1	1	1	1	11	4	3	2	1	32	17	100

Comparison of Manual Occupation Codes to Auto-Coded Occupation (70% Confidence)

Manual Occupation Codes	NIOCCS Generated Occupation Codes																							
	031	071	263	313	394	411	422	423	476	540	562	605	620	626	720	772	901	905	906	910	913	990	blank	Total
031	1																							1
263			1																					1
360																							2	2
365																							1	1
394					1																			1
402																							2	2
411						3																	1	4
414																							1	1
422							1																	1
423				1				1															1	3
425																							2	2
461																							1	1
476									1															1
540										1														1
570	1																						1	2
586																							1	1
605												1												1
626														3										3
720															1									1
739		1																						1
772																1								1
775																							1	1
780																							1	1
781																							1	1
901																	11							11
905																		4						4
906																			3				1	4
910																				2				2
913																					1			1
962											1													1
990	1												1									32	8	42
Total	3	1	1	1	1	3	1	1	1	1	1	1	1	3	1	1	11	4	3	2	1	32	25	100

Comparison of Manual Occupation Codes to Auto-Coded Occupation (90% Confidence)

Manual Occupation Codes		NIOCCS Generated Occupation Codes															
		031	263	313	411	422	423	476	626	720	901	905	906	910	990	blank	Total
	031	1															1
	263		1														1
	360															2	2
	365															1	1
	394															1	1
	402															2	2
	411				3											1	4
	414															1	1
	422					1											1
	423			1			1									1	3
	425															2	2
	461															1	1
	476							1									1
	540															1	1
	570															2	2
	586															1	1
	605															1	1
	626								3								3
	720									1							1
	739															1	1
	772															1	1
	775															1	1
	780															1	1
	781															1	1
	901										11						11
	905											4					4
	906												3			1	4
	910													2			2
	913															1	1
	962															1	1
	990														32	10	42
	Total	1	1	1	3	1	1	1	3	1	11	4	3	2	32	35	100

REFERENCES

(1) Halperin W, Howard J, Centers for Disease Control and Prevention (CDC). Occupational epidemiology and the National Institute for Occupational Safety and Health. MMWR Surveill Summ 2011 Oct 7;60 Suppl 4:97-103.

(2) Checkoway H, Pearce N, Kriebel D. Research Methods in Occupational Epidemiology. Monographs in Epidemiology and Biostatistics. 2nd ed. New York: Oxford University Press; 2004.

(3) About OSHA. Available at: http://www.osha.gov/about.html. Accessed May 29, 2012.

(4) About NIOSH. 2011; Available at: http://www.cdc.gov/niosh/about.html. Accessed May 29, 2012.

(5) Schulte PA. Characterizing the Burden of Occupational Injury and Disease. Journal of Occupational and Environmental Medicine 2005;47(6).

(6) Leigh JP. Economic burden of occupational injury and illness in the United States. Milbank Q 2011 Dec;89(4):728-772.

(7) 91st Congress S2. Occupational Safety and Health Act of 1970 : SEC. 2. Congressional Findings and Purpose (29 USC 651). 2004; Available at: http://www.osha.gov/pls/oshaweb/owadisp.show_document?p_table=OSHACT&p_id=2743. Accessed May 29, 2012.

(8) Lurio J, Morrison FP, Pichardo M, Berg R, Buck MD, Wu W, et al. Using electronic health record alerts to provide public health situational awareness to clinicians. Journal of the American Medical Informatics Association 2010 March 01;17(2):217-219.

(9) Shapiro JS. Evaluating public health uses of health information exchange. J Biomed Inform 2007 Dec;40(6 Suppl):S46-9.

(10) Health Resources and Services Administration (HRSA). What is health information exchange? 2012; Available at: http://www.hrsa.gov/healthit/toolbox/RuralHealthITtoolbox/Collaboration/whatishie.html. Accessed January 15, 2013.

(11) Vest JR, Gamm LD. Health information exchange: persistent challenges and new strategies. Journal of the American Medical Informatics Association 2010 May 01;17(3):288-294.

(12) NORA AgFF Sector Council. National Agriculture, Forestry, and Fishing Agenda. 2008.

(13) Committee on the Respiratory Protection Curriculum for Occupational Health Nursing Programs, Institute of Medicine. Incorporating Occupational Information in Electronic Health Records: Letter Report. : The National Academies Press; 2011.

(14) Bureau of Labor Statistics, U.S. Department of Labor. Occupational Outlook Handbook, 2012-13 Edition, Agricultural Workers. 2012; Available at: http://www.bls.gov/ooh/farming-fishing-and-forestry/agricultural-workers.htm. Accessed April 08, 2012.

(15) Larson AC. Migrant and Seasonal Farmworker Enumeration Profiles Study. 2000; Available at: http://www.ncfh.org/enumeration/PDF10%20Texas.pdf. Accessed April 10, 2010.

(16) Carroll D, Samardick RM, Bernard S, Gabbard S, Hernandez T. Findings from the National Agricultural Workers Survey (2001-2002). A Demographic and Employment Profile of United States Farm Workers. 2005; Available at: http://www.doleta.gov/agworker/report9/naws_rpt9.pdf. Accessed April 07, 2012.

(17) US Department of Health and Human Services. Health Resources and Services Administration (HRSA) Primary Care: The Health Center Program "Special Populations". Available at: http://bphc.hrsa.gov/about/specialpopulations/index.html. Accessed April 11, 2012.

(18) Brower MA, Earle-Richardson GB, May JJ, Jenkins PL. Occupational injury and treatment patterns of migrant and seasonal farmworkers. J Agromedicine 2009;14(2):172-178.

(19) Centers for Medicare and Medicaid Services. The Official Web Site for the Medicare and Medicaid Electronic Health Records (EHR) Incentive Programs. 2012; Available at: https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/. Accessed April 10, 2012.

(20) Centers for Medicare and Medicaid Services. Electronic Health Records. 2012; Available at: https://www.cms.gov/Medicare/E-Health/EHealthRecords/. Accessed April 8, 2012.

(21) Centers for Medicare and Medicaid Services. CMS EHR Meaningful Use Overview. 2012; Available at: https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/Meaningful_Use.html. Accessed April 8, 2012.

(22) National Institute for Occupational Safety and Health Division of Surveillance, Hazard Evaluations, and Field Studies. NIOSH Industry & Occupation Computerized Coding System (NIOCCS). December 20, 2011; Available at: http://www.cdc.gov/niosh/topics/coding/software.html, 2012.

(23) National Center for Farmworker Health (NCFH). 2011; Available at: http://www.ncfh.org/, 2012.

(24) Garcia D, Hopewell J, Liebman AK, Mountain K. The migrant clinicians network: connecting practice to need and patients to care. J Agromedicine 2012 Jan;17(1):5-14.

(25) Cooper SP, Shipp EM, Felknor SA, del Junco DJ, Hendrickson E, Heyer N, et al. Community-Academic Linkages to Build a National Network for Farmworker Research. National Institute of Environmental Health Sciences (NIEHS) 2010 2013;RFA-OD-09-010(999,999).

(26) Trotter F, Uhlman D. Meaningful Use and Beyond. 1st ed. Sebastopol, California: O'Reilly Media, Inc.; 2011.

(27) Julie Graves-Moy. Analysis of policy issues surrounding the electronic medical record. Houston, Texas: University of Texas School of Public Health; 2011.

(28) Centers for Disease Control and Prevention, National Center for Health Statistics. International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM). October 4, 2012; Available at: http://www.cdc.gov/nchs/icd/icd9cm.htm. Accessed October 8, 2012.

(29) About Health Level Seven (HL7) International. 2012; Available at: http://www.hl7.org/about/index.cfm?ref=footer. Accessed December 8, 2012.

(30) Office of the National Coordinator (ONC) for Health Information Technology. The Certified HIT Product List (CHPL). June 26, 2012; Available at: http://oncchpl.force.com/ehrcert?q=CHPL. Accessed October 8, 2012.

(31) Ancker JS, Edwards AM, Miller MC, Kaushal R. Consumer perceptions of electronic health information exchange. Am J Prev Med 2012 Jul;43(1):76-80.

(32) Patel VN, Abramson E, Edwards AM, Cheung MA, Dhopeshwarkar RV, Kaushal R. Consumer attitudes toward personal health records in a beacon community. Am J Manag Care 2011 Apr;17(4):e104-20.

(33) Ancker JS, Silver M, Miller MC, Kaushal R. Consumer experience with and attitudes toward health information technology: a nationwide survey. J Am Med Inform Assoc 2012 Jul 30.

(34) Ancker JS, Barron Y, Rockoff ML, Hauser D, Pichardo M, Szerencsy A, et al. Use of an electronic patient portal among disadvantaged populations. J Gen Intern Med 2011 Oct;26(10):1117-1123.

(35) Kukafka R, Ancker JS, Chan C, Chelico J, Khan S, Mortoti S, et al. Redesigning electronic health record systems to support public health. J Biomed Inform 2007 Aug;40(4):398-409.

(36) Calderwood MS, Platt R, Hou X, Malenfant J, Haney G, Kruskal B, et al. Real-time surveillance for tuberculosis using electronic health record data from an ambulatory practice in eastern Massachusetts. Public Health Rep 2010 Nov-Dec;125(6):843-850.

(37) Hripcsak G, Soulakis ND, Li L, Morrison FP, Lai AM, Friedman C, et al. Syndromic Surveillance Using Ambulatory Electronic Health Records. Journal of the American Medical Informatics Association 2009 May 01;16(3):354-361.

(38) International Society for Disease Surveillance (ISDS). Final Recommendation: Core Processes and EHR Requirements for Public Health Syndromic Surveillance. 2011; Available at: http://www.syndromic.org/uploads/files/ISDSRecommendation_FINAL.pdf. Accessed October 2, 2012.

(39) Khurshid A, Diana ML, Luce SD. Health information exchange: metrics to address quality of care and return on investment. Perspect Health Inf Manag 2012;9:1e.

(40) Pevnick JM, Claver M, Dobalian A, Asch SM, Stutman HR, Tomines A, et al. Provider stakeholders' perceived benefit from a nascent health information exchange: a qualitative analysis. J Med Syst 2012 Apr;36(2):601-613.

(41) Howard J, Clark EC, Friedman A, Crosson JC, Pellerano M, Crabtree BF, et al. Electronic Health Record Impact on Work Burden in Small, Unaffiliated, Community-Based Primary Care Practices. J Gen Intern Med 2012 Aug 28.

(42) Wiljer D, Urowitz S, Apatu E, DeLenardo C, Eysenbach G, Harth T, et al. Patient accessible electronic health records: exploring recommendations for successful implementation strategies. J Med Internet Res 2008 Oct 31;10(4):e34.

(43) Taylor JA, Frey LT. The Need for Industry and Occupation Standards in Hospital Discharge Data. J Occup Environ Med 2013 Apr 24.

(44) Friedman L, Krupczak C, Brandt-Rauf S, Forst L. Occupational amputations in Illinois 2000-2007: BLS vs. data linkage of trauma registry, hospital discharge, workers compensation databases and OSHA citations. Injury 2012 Feb 23.

(45) Texas Department of State Health Services (TXDSHS). Pesticide Exposure Surveillance in Texas. 2012; Available at: http://www.dshs.state.tx.us/epitox/Pesticide-Exposure/. Accessed 10/8, 2012.

(46) Public Health Surveillance and Informatics Program Office, Centers for Disease Control and Prevention (CDC). National Notifiable Disease Surveillance System (NNDSS). 2012; Available at: http://wwwn.cdc.gov/nndss/document/2012_Case%20Definitions.pdf. Accessed September 15, 2012.

(47) 29 USC Chapter 20 - Migrant and Seasonal Agricultural Worker Protection.

(48) US Department of Health and Human Services. Health Resources and Services Administration (HRSA) Primary Care: The Health Center Program "What is a health center?". Available at: http://bphc.hrsa.gov/about/. Accessed April 11, 2012.

(49) National Center for Farmworker Health (NCFH). About Community and Migrant Health Centers. Available at: http://www.ncfh.org/?sid=37. Accessed October 5, 2012.

(50) Health Resources and Services Administration (HRSA) Primary Care: The Health Center Program. Electronic Health Record (EHR) Information 2011 National Data: 1128 Grantees. Available at: http://bphc.hrsa.gov/uds/view.aspx?q=rehr&year=2011&state=. Accessed September 28, 2012.

(51) National Advisory Council on Migrant Health (NACMH). NACMH's Recommendations to the Secretary of HHS. 2012; Available at: http://bphc.hrsa.gov/about/nacmh/recommendations/index.html. Accessed October 10, 2012.

(52) Earle-Richardson GB, Brower MA, Jones AM, May JJ, Jenkins PL. Estimating the occupational morbidity for migrant and seasonal farmworkers in New York state: a comparison of two methods. Ann Epidemiol 2008 Jan;18(1):1-7.

(53) U.S. Department of Labor. National Agricultural Workers Survey (NAWS). 2011; Available at: http://www.doleta.gov/agworker/naws.cfm. Accessed December 01, 2012.

(54) Hansen E, Donohoe M. Health issues of migrant and seasonal farmworkers. J Health Care Poor Underserved 2003 May;14(2):153-164.

(55) Colt JS, Stallones L, Cameron LL, Dosemeci M, Zahm SH. Proportionate mortality among US migrant and seasonal farmworkers in twenty-four states. Am J Ind Med 2001;40(5):604-611.

(56) Dever A. Migrant Health Status: Profile of a Population With Complex Health Problems. Austin, Texas: Migrant Clinicians Network Monograph Series; 1991.

(57) Villarejo D, McCurdy SA, Bade B, Samuels S, Lighthall D, Williams D,3rd. The health of California's immigrant hired farmworkers. Am J Ind Med 2010 Apr;53(4):387-397.

(58) Spurzem JR, Romberger DJ, Von Essen SG. Agricultural lung disease. Clin Chest Med 2002 Dec;23(4):795-810.

(59) Mills PK, Dodge J, Yang R. Cancer in migrant and seasonal hired farm workers. J Agromedicine 2009;14(2):185-191.

(60) Shah DJ, Shipp EM, Cooper SP, Huber JC, del Junco DJ, Rene AA, et al. Hand problems in migrant farmworkers. Journal of Agricultural Safety & Health 2009 Apr;15(2):157-169.

(61) Irby CE, Yentzer BA, Vallejos QM, Arcury TA, Quandt SA, Feldman SR. The prevalence and possible causes of contact dermatitis in farmworkers. Int J Dermatol 2009 Nov;48(11):1166-1170.

(62) Weigel MM, Armijos RX. Exploratory study of the occupational health and health-seeking of migrant and seasonal farmworkers on the u.s.-Mexico border. J Immigr Minor Health 2012 Aug;14(4):648-656.

(63) Emmi KE, Jurkowski JM, Codru N, Bell EM, Kacica MA, Carter TP. Assessing the health of migrant and seasonal farmworkers in New York State: statewide data 2003-2005. J Health Care Poor Underserved 2010 May;21(2):448-463.

(64) Anthony M, Williams JM, Avery AM. Health needs of migrant and seasonal farmworkers. J Community Health Nurs 2008 Jul-Sep;25(3):153-160.

(65) Villarejo D, McCurdy SA. The California Agricultural Workers Health Survey. J Agric Saf Health 2008 Apr;14(2):135-146.

(66) Wang S, Myers JR, Layne LA. Injuries to hired crop workers in the United States: a descriptive analysis of a national probability survey. Am J Ind Med 2011 Oct;54(10):734-747.

(67) Ward LS. Preliminary tests of an ecological model of Hispanic farmworker health. Public Health Nurs 2007 Nov-Dec;24(6):554-564.

(68) National Center for Farmworker Health (NCFH), 22nd Annual Midwest Stream Forum. Occupational Health Findings from the National Agricultural Worker Survey (NAWS); November 8, 2012; Buda, Texas: National Center for Farmworker Health (NCFH); 2012.

(69) Cooper SP, Burau KE, Frankowski R, Shipp EM, Del Junco DJ, Whitworth RE, et al. A cohort study of injuries in migrant farm worker families in South Texas. Ann Epidemiol 2006 Apr;16(4):313-320.

(70) Galewitz P. Fear keeps migrant workers from getting health care. USA Today; Kaiser Health News 2012 6/6/2012;Nation.

(71) Anthony MJ, Martin EG, Avery AM, Williams JM. Self Care and Health-Seeking Behavior of Migrant Farmworkers. Journal of Immigrant and Minority Health 2010 Oct 2010;12(5):634-9.

(72) Flocks J. Pesticide policy and farmworker health. Rev Environ Health 2009 Oct-Dec;24(4):327-332.

(73) Grzywacz JG, Alterman T, Muntaner C, Gabbard S, Nakamoto J, Carroll DJ. Measuring job characteristics and mental health among Latino farmworkers: results from cognitive testing. J Immigr Minor Health 2009 Apr;11(2):131-138.

(74) Cooper SP, Burau K, Hanis C, Henry J, MacNaughton N, Robison T, et al. Tracing migrant farmworkers in Starr County, Texas. Am J Ind Med 2001 Nov;40(5):586-591.

(75) United States Department of Health and Human Services (HHS). Health People 2020 Objective Topic Areas. 2010; Available at: http://www.healthypeople.gov/2020/topicsobjectives2020/pdfs/HP2020objectives.pdf . Accessed December 10, 2012.

(76) Sears JM, Bowman SM, Adams D, Silverstein BA. Occupational injury surveillance using the Washington State Trauma Registry. J Occup Environ Med 2011 Nov;53(11):1243-1250.

(77) Lighthall D. The poor health of farm workers. West J Med 2001 Oct;175(4):223-224.

(78) Hiott AE, Grzywacz JG, Davis SW, Quandt SA, Arcury TA. Migrant Farmworker Stress: Mental Health Implications. The Journal of Rural Health 2008;24(1):32-39.

(79) Schumacher P, editor. Western States Occupational Network (WESTON) Third Annual Meeting NIOSH/CSTE/MAP-ERC. Industry and occupation coding – updates and methods; September 23, 2010; 2010.

(80) Ruder AM, Waters MA, Carreon T, Butler MA, Calvert GM, Davis-King KE, et al. The Upper Midwest Health Study: Industry and occupation of glioma cases and controls. Am J Ind Med 2012 Jun 19.

(81) Cherniack M, Henning R, Merchant JA, Punnett L, Sorensen GR, Wagner G. Statement on national worklife priorities. Am J Ind Med 2011 Jan;54(1):10-20.

(82) United States Census Bureau, Housing and Household Economic Statistics Division. 2000 Industry and Occupation Classification System. 2012; Available at: http://www.census.gov/hhes/www/ioindex/search.html. Accessed January 15, 2013.

(83) Centers for Disease Control and Prevention (CDC). Vital signs: awareness and treatment of uncontrolled hypertension among adults--United States, 2003-2010. MMWR Morb Mortal Wkly Rep 2012 Sep 7;61:703-709.

(84) Di Legge S, Koch G, Diomedi M, Stanzione P, Sallustio F. Stroke Prevention: Managing Modifiable Risk Factors. Stroke Research and Treatment 2012;2012:15.

(85) Centers for Disease Control and Prevention (CDC), National Center for Health Statistics. Faststats-Leading causes of death. 2012; Available at: http://www.cdc.gov/nchs/fastats/lcod.htm. Accessed November 30, 2012.

(86) U.S. Department of Health and Human Services (HHS), National Institutes of Health (NIH), National Heart, Lung, and Blood Institute, National High Blood Pressure Education Program. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. 2004 August 2004;7.

(87) Luque JS, Reyes-Ortiz C, Marella P, Bowers A, Panchal V, Anderson L, et al. Mobile farm clinic outreach to address health conditions among Latino migrant farmworkers in Georgia. J Agromedicine 2012 Oct;17(4):386-397.

(88) Ostchega Y, Yoon SS, Hughes J, Louis T. Hypertension awareness, treatment, and control--continued disparities in adults: United States, 2005-2006. NCHS Data Brief 2008 Jan(3):1-8.

(89) Olives C, Myerson R, Mokdad AH, Murray CJ, Lim SS. Prevalence, awareness, treatment, and control of hypertension in United States counties, 2001-2009. PLoS One 2013;8(4):e60308.

(90) Rehman SU, Hutchison FN, Hendrix K, Okonofua EC, Egan BM. Ethnic differences in blood pressure control among men at Veterans Affairs clinics and other health care sites. Arch Intern Med 2005 May 9;165(9):1041-1047.

(91) Romero CX, Romero TE, Shlay JC, Ogden LG, Dabelea D. Changing trends in the prevalence and disparities of obesity and other cardiovascular disease risk factors in three racial/ethnic groups of USA adults. Adv Prev Med 2012;2012:172423.

(92) Kannel WB. Risk stratification in hypertension: new insights from the Framingham study. American Journal of Hypertension 2000 1;13(1, Supplement 1):S3-S10.

(93) Ferdinand KC. Managing cardiovascular risk in minority patients. J Natl Med Assoc 2005 Apr;97(4):459-466.

(94) Sechrest L. Validity of measures is no simple matter. Health Serv Res 2005 Oct;40(5 Pt 2):1584-1604.

(95) Calvert GM, Karnik J, Mehler L, Beckman J, Morrissey B, Sievert J, et al. Acute pesticide poisoning among agricultural workers in the United States, 1998-2005. Am J Ind Med 2008 Dec;51(12):883-898.

(96) Casarett LJ, Doull J. Toxic Effects of Pesticides. In: Klaassen CD, editor. Casarett and Doull's toxicology: The basic science of poisons. 7th ed. New York: McGraw-Hill Medical; 2008. p. 883.

(97) United States Environmental Protection Agency (EPA). Worker Protection Standards for Agricultural Pesticides. 2012; Available at: http://www.epa.gov/agriculture/twor.html. Accessed December 10, 2012.

(98) Auf der Heide E. Case Studies in Envrionmental Medicine- Cholinesterase Inhibitors Including Pesticides and Chemical Warfare Nerve Agents. 2007.

(99) Chen Y. Organophosphate-induced brain damage: Mechanisms, neuropsychiatric and neurological consequences, and potential therapeutic strategies. Neurotoxicology 2012 6;33(3):391-400.

(100) Calvert GM. Case Definition for Acute Pesticide-Related Illness and Injury Cases Reportable to the National Public Health Surveillance System. 2005; Available at: http://wwwn.cdc.gov/niosh-survapps/sensor/CaseDefinition.pdf. Accessed December 1, 2012.

(101) Shipp EM, Cooper SP, Donnelly KC, Nalbone JT. A feasibility study of the use of dust wipe samples to assess pesticide exposures in migrant farmworker families. J Agromed 2006;11(3-4):83-93.

(102) Blair A, Thomas K, Coble J, Sandler DP, Hines CJ, Lynch CF, et al. Impact of pesticide exposure misclassification on estimates of relative risks in the Agricultural Health Study. Occup Environ Med 2011 Jul;68(7):537-541.

(103) United States Environmental Protection Agency (EPA). Worker Protection Standards for Agricultural Pesticides. 2012; Available at: http://www.epa.gov/agriculture/twor.html. Accessed December 1, 2012.

(104) Arcury TA, Quandt SA. Living and working safely: challenges for migrant and seasonal farmworkers. N C Med J 2011 Nov-Dec;72(6):466-470.

(105) Texas Department of State Health Services (TXDSHS). Who is responsible for reporting occupational conditions? 2013; Available at: http://www.dshs.state.tx.us/epitox/Pesticide-Exposure/#reporting. Accessed January 4, 2013.

(106) Kasner EJ, Keralis JM, Mehler L, Beckman J, Bonnar-Prado J, Lee SJ, et al. Gender differences in acute pesticide-related illnesses and injuries among farmworkers in the United States, 1998-2007. Am J Ind Med 2012 Jul;55(7):571-583.

(107) Ogut S, Gultekin F, Kisioglu AN, Kucukoner E. Oxidative stress in the blood of farm workers following intensive pesticide exposure. Toxicol Ind Health 2011 Oct;27(9):820-825.

(108) Calvert GM, Higgins SA. Using surveillance data to promote occupational health and safety policies and practice at the state level: A case study. Am J Ind Med 2010;53(2):188-193.

(109) Levesque DL, Arif AA, Shen J. Effectiveness of pesticide safety training and knowledge about pesticide exposure among Hispanic farmworkers. J Occup Environ Med 2012 Dec;54(12):1550-1556.

(110) Levesque DL, Arif AA, Shen J. Association between Workplace and Housing Conditions and Use of Pesticide Safety Practices and Personal Protective Equipment among North Carolina Farmworkers in 2010. Int J Occup Environ Med 2012 Apr;3(2):53-67.

(111) Shipp EM, Cooper SP, del Junco DJ, Bolin JN, Whitworth RE, Cooper CJ. Pesticide safety training among farmworker adolescents from Starr County, Texas. Journal of Agricultural Safety & Health 2007 Jul;13(3):311-321.

(112) Shipp EM, Cooper SP, Burau KD, Bolin JN. Pesticide safety training and access to field sanitation among migrant farmworker mothers from Starr County, Texas. Journal of Agricultural Safety & Health 2005 Feb;11(1):51-60.

(113) Oliveira Pasiani J, Torres P, Roniery Silva J, Diniz BZ, Dutra Caldas E. Knowledge, attitudes, practices and biomonitoring of farmers and residents exposed to pesticides in Brazil. Int J Environ Res Public Health 2012 Aug 24;9(9):3051-3068.

(114) Spears CR, Summers PY, Spencer KM, Arcury TA. Informal Occupational Safety Information Exchange Among Latino Migrant and Seasonal Farmworkers. J Agromed 2012 10/01; 2012/12;17(4):415-420.

(115) Salud Family Health Centers. 2009; Available at: http://www.saludclinic.org/. Accessed January 2, 2012.

(116) StataCorp. *Stata Statistical Software*. 2009;11.

(117) Centers for Disease Control and Prevention (CDC), National Center for Health Statistics. Instruction Manual Part 19- Industry and Occupation Coding for Death Certificates. 2003.

(118) Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. Biometrics 1977 Mar.;33(1):159-174.

(119) Nigg HN, Knaak JB. Blood cholinesterases as human biomarkers of organophosphorus pesticide exposure. Rev Environ Contam Toxicol 2000;163:29-111.

(120) Keifer M, Sheridan C, Liebman A. Cholinesterase Testing Protocols. Migrant Clinicians Network (Clinical Tools: Pesticides).

(121) Farmworker Justice. Workers' Compensation. 2013; Available at: http://farmworkerjustice.org/content/workers-compensation. Accessed April 19, 2013.

(122) Rodríguez AAA, Rodilla JMR, Delclos J. Documenting occupation in the medical chart of a public hospital system: 2006-2010. Archivos de prevención de riesgos laborales 2012;15(3):124.

(123) Ditolvi Vera G, Benavides Márquez F, Armengol O, Barrionuevo-Rosas L. Cumplimentación de la ocupación en las historias clínicas de la Atención Primaria: 1992 2007. Atención Primaria 2010;42:486-487.

(124) Szklo M, Nieto FJ editors. Epidemiology Beyond the Basics. 2nd ed. Sudbury, MA: Jones and Bartlett Publishers; 2007.

(125) Buntin MB, Burke MF, Hoaglin MC, Blumenthal D. The benefits of health information technology: a review of the recent literature shows predominantly positive results. Health Aff (Millwood) 2011 Mar;30(3):464-471.