Phases

PHASE 1: DATA MATCHING

- Can the MAAP-NS dataset and HDNS datasets be used to validate and supplement each other where each contains similar variables

From MAAP-NS	From HDNS	Gold standard	How they will be used
Provider **age**	Licensed provider registry	MAAP-NS	HDNS data will serve to fill in missing values for MAAP-NS data
Provider **sex**	Licensed provider registry	MAAP-NS	HDNS data will serve to fill in missing values for MAAP-NS data
PracPop (Provider's estimate of practice size)	The number of unique patients each provider sees in each of the three years of the study	The count from HDNS datasets will be considered to be accurate	We will determine whether providers typically over- or under-estimate the size of their practices and by how much.
ERwork (y/n) Helps define scope of practice	MSI Physician Billings Location=HOSP; Hospunit=EMCC Captures visits to ED	MAAP-NS responses will be accepted first; HDNS data may fill in missing values	A provider's ER billings would validate the response to ERwork (y/n) from the MAAP-NS survey and possibly fill-in missing information
HomeVisit (y/n) Helps define scope of practice and high responsibility provider	MED (MSI Physician Billings) Location=HOME; HMHC	HDNS responses will be accepted first; MAAP-NS data may fill in missing values	A provider's billings for home visits would validate the response to Home Visit (y/n) from the MAAP-NS survey and possibly fill-in missing information
InPt (y/n) informs the designation of "high" vs "low" responsibility provider	MED (MSI Physician Billings) Location=HOSP; Hospunit=INPT Captures visits to patients in hospitals	HDNS responses will be accepted first; MAAP-NS data may fill in missing values	A provider's billings for inpatient visits would validate the response to InPt (y/n) from the MAAP-NS survey and possibly fill-in missing information

- A report of the results of this comparison will be delivered to the PI for review before further work is done.

PHASE 2: DESCRIPTIVE SUMMARY AND STATISTICS

The following tables created in Deliverables session will be delivered to PI as second phase study. MSSU will check normality for each comparison, distribution (central tendency, distribution and range) of data will be reported to PI. The first analyses will be run on the data in crude format; subsequent analyses of certain variables will need to include adjustment for such things as age, sex and location at the direction of the PI.

- Table 1.1: The relationship between means of patients ED visits and means of patients' primary care visits during regular office hours (8am-5pm weekdays) and after-hours (5:01pm - 7:59am weekdays and anytime weekends) and altogether; and whether the patient's main provider was offering appointments at the time of the ED visit (from survey data).

- Table1.1.1: The relationship between the number of non-admitted ED visits and the number of visits to a primary care provider over 3 years and individually by each of these 3 years.

- Table 1.1.2: The relationship between at least one admission from the ED in the third year or not and the number of primary care visits in the previous 2 years; also the relationship between at least 1 admission from ED and the rate of non-admitted ED visits in the previous 2 years.

- Table 1.1.3: The relationship between the number of primary care providers seen by the patient. One we look at this distribution we will examine the relationship between number of primary care providers seen, number of visits to all providers and ED use over 1 year, 2 years, 3 years.

- Table2: Mean number of ED visits (FY2012-2014) during regular office hours (8am-5pm weekdays) and after-hours (5:01pm - 7:59am weekdays; anytime Saturdays and Sundays) and altogether

- Table3.1: Mean number of ED Uses (FY2012-2014)

- Table3.2: Compare number of ED Uses and urgent and non-urgent wait times

- Table3.3: Compare urgent and non-urgent wait times and one or multiples primary care providers for all patients

- Table 3.3.1: Compare urgent and non-urgent wait times and one or multiples primary care providers for patients with mental illness

- Table 3.3.2: Compare urgent and non-urgent wait times and one or multiples primary care providers for patients with specific chronic physical conditions

- Table 4: Models of care vs. patient ED use (total # of ED visits made by a provider's patients divided by the number of patients of the provider). This will be numeric & continuous, may need to be categorized once frequency &

distribution are known. Details of the classifications for "models of care" variables are provided. Additional analysis may be done and will be determined once preliminary results for ED use are known. This is an example of where you would have to adjust for any gaps in service by providers or gaps for patients.

- Table 4.1: There are an initial 7 classifications of models of care (see Table 4 page 39). ED use (total # of ED visits made by a provider's patients divided by the number of patients of the provider) will be compared between models of care. This will be numeric & continuous, and may need to be categorized once freq & distribution are known. This is an example of where you would have to adjust for any gaps in service by providers or gaps for patients.

Additional details of the "models of care" variables and how this analysis will be done will be determined once preliminary results for ED use are known.

- Table 5: The distribution of "Provider Responsibility"
- Table 6.1: The frequency of family physician visits
- Table 6.2: The means of family physician visits
- Table 7.1: The rate of screening procedures
- Table 8.1: Scope of practice related to patient ED use and continuity

PHASE 3: STATISTICS AND REGRESSION ANALYSIS

NOTE: PHASE 3 & PHASE 4 TABLES AND METHODOLOGIES ARE INCLUDED HERE AS A STARTING POINT FOR THE REGRESSION ANALYSES AND ANALYSES INVOLVING MODELS OF CARE. THESE FIRST RUNS OF REGRESSION ANALYSES CAN BE COMPLETED AS DESCRIBED WITHOUT WAITNG FOR A SUBSEQUENT CHARTER. ANY ADDITIONAL DEFINITIONS AND CATEGORIZATIONS WILL BE PROPOPSED AFTER THIS WORK.

DEFINITIONS FOR the classifications of models of care, high to Low responsibility providers are provided below.

The type of regression analyses employed (e.g., negative binomial, Poisson, logistic regression) will be determined once the data has been reviewed. (Phase 2 and 3). The method of following analysis will be determined by the PI.

1. Objective 1.2) : the relationship between ED visits and primary care visits influenced by patient characteristics, by practice & provider characteristics derived from the survey data (age, sex, hours of availability; wait times for urgent and non-urgent care)

2. Objective (2): The relationship between emergency department visits among patients and provide afterhours care of their NP/family physician among the patient's age; sex; presence of chronic illness and mental illness.

3. Objective (3): Wait times for urgent and non-urgent care are associated with ED use, additional provider use.

4. Objective(6): Do patterns of NP/family physician utilization (how many unique providers and number of visits) differ by models of primary care, geographic location, and accessibility or provider characteristics

    - Outcome variables: The frequency of NP/family physician visits
    - Univariate relationship: adjusted by models of primary care; geographic location (both management zone and also urban/rural categories); after hours accessibility; physician characteristics (age, sex, country of training). [Western/not; Canada/not] as individual predictors. There are 2 ways of looking at location and doctors training. So this would have to have 4 separate models run with the different combinations.
    - Table 6.3: The univariate regression of NP/family physician visits

PHASE 4: MULTIPLE REGRESSION ANALYSES

Note, for all multiple regression: The first analyses will be run on the data in crude format; subsequent analyses of certain variables will need to include adjustment for such things as age, sex and location at the direction of the PI.

1. The subject (6): Do patterns of family physician utilization (how many unique providers and number of visits) differ by models of primary care, geographic location, and accessibility or provider characteristics

Methodology for Objective 1

1.1 Correlation of frequency of primary care visits {visits in office or home; not LTC or ED) per year with frequency of ED visits per year (Pearson's r for normally distributed data; Spearman's rho if the data is skewed).

1.1.1 Correlation as in **1.1**, using the subset of patients whose ED visits did not result in admission to hospital.

1.1.2 Patients who have made ED visits in 2013-2014 will be categorized according to whether their ED visit resulted in hospital admission or not. The frequency of each groups' ED visits and the frequency of their primary care visits during the previous two years will be compared between these two groups of patients using t-tests for normally distributed data or Wilcoxon Rank Sum tests for non-normal data.

1.2 Multiple regression analysis predicting frequency of ED visits where the variables of patient age and sex and the characteristics of the pre-determined usual provider (age, sex, hours of availability; wait times for urgent and non-urgent care) will be included in the model.

1.3 The number of unique primary care providers (family physicians and nurse practitioners) seen by each patient during each year will be computed and the relationship of this measure to the frequency of ED visits will be examined using correlations. Patients will be categorized according to number of primary care providers (depending on the distribution of the numbers of providers) seen in one year, two years and three years. The number of ED visits by each group of primary care consumers (e.g., single versus multiple) will be compared for each of these time frames using t-tests for normally distributed data or Wilcoxon Rank Sum tests for non-normal data.

Methodology for Objective 2

Family physicians can be categorized as providing after hours care (yes/no) from the MAAP-NS survey data. There are two types of "after hours" categories: 1) working before 7 am or working after 6 pm as these times correspond to incentive funding; 2) before 9 am or after 5 pm as these times correspond to regular business hours.

Chi-square analysis will compare the proportions of patients of providers in each category who accessed EDs. This comparison will be repeated using the subset of patients with mental illness codes and again for the subset of patients with any of three chronic illnesses (diabetes, COPD or cardiovascular disease) as defined by the Canadian Chronic Disease Surveillance System. Next, the contribution of after-hours care category to the ED visits will be entered into a logistic regression model. The outcome will be "ED visit (yes/no)." Other predictors included in the model will be patient's age; sex; presence of chronic illness: COPD (y/n), diabetes (y/n), cardiovascular disease (y/n); presence of mental illness (y/n), and morbidity index.

Methodology for Objective 3

The wait time variables (from the survey data) are both continuous (number of days) and categorical. The relationships between the wait times in days and the frequency of ED use will be examined with correlations (Pearson's r or Spearman's rho depending on normality of data). The frequency of ED use among the categories of urgent [i.e., same day, next day, 2-5 days, >5 days] and non-urgent [i.e., ≤1 day, 2-5 days, 6-10 days, >10 days] wait times will be examined using chi-square analyses. These analyses will be repeated with the subsets of patients having mental illness, heart disease, COPD, and diabetes.

Patients will be categorized as having typically one or multiple primary care providers. The continuous measures of urgent and non-urgent wait times associated with the usual provider

will be compared between these two groups of patients using t-tests or Wilcoxon Rank Sum tests, depending on normality of the data.

To further delineate the contribution of wait times to frequency of ED use, a linear regression model will be constructed to predict frequency of ED use with the wait time variables as predictors as well as patient, provider and practice characteristics [e.g., patient and provider sex and age, size and model of practice, and hours of availability of provider and morbidity index].

Methodology for Objective 4

Classifications of Models of primary care will be developed from the MAAP-NS survey (size, location & make-up of practice). The preliminary 7 classifications are described below (page 39)

Patients ED use will be a simple count of visits to ED for each fiscal year.

Practice burden of chronic disease management will be expressed as the number of CDM codes each primary health care provider submits divided by the total number of individual patients seen who have each of these diseases (Diabetes, COPD, Ischemic heart disease by that provider in each fiscal year. Rates of CDM provision will be compared among the models of care with chi-square analysis.

Methodology for Objective 5

The issue of lower to higher responsibility will be reflected in a composite measure which will include:

- rate of referrals: the number of referrals made divided by the total number of unique patients seen by that provider over the 3-year study period
- provision of out-of-office visits, looking at patient homes, and in-hospital separately
- and, provision of the following preventative measures by the proportion of patients for each year for each physician in the study (other eligibility described below):
    - Proportion of one or more PAP tests for females aged 21-69 inclusive over three years and over five years
    - number of vaccines for tuberculosis divided by the number of patients per year per provider
    - number of vaccinations for measles/mumps/rubella (as a proportion of patients under the age of 5) per provider
    - number of vaccines for Adacel or Boostrix (diphtheria, pertussis and tetanus), Adacel-Polio, divided by the number of patients 23 years of age and older per provider, going back over 5 years of data. (they are to be given once every 5+ years)

- number of vaccines for pneumococcal pneumonia divided by the number of patients >=65 seen. Looking at 5 years of data, we want to know what proportion of patients 65 and over have received this vaccination divided by all the patients in that age group, per year, per provider
- number of influenza vaccinations divided by the number of patients per year for each provider
- number of varicella vaccinations divided by the number of patients aged 60 and over per year, per provider

The proportion of high responsibility providers will be compared among the models of care and ED use with Chi-square analysis. These analyses will be repeated with the subset of patients with mental illness and those with the specific chronic diseases.

Methodology for Objective 6

Frequency and distribution of family physician/nurse practitioner visits will be determined for each patient linked to a provider from the MAAP-NS survey. Depending on the normality of the distribution of these visits, they will be treated as either a continuous or categorical variable (wherein the frequencies will be categorized in terms of median and above versus below the median). The frequencies (or categories) will be presented in terms of the univariate relationships to the variables: models of primary care; geographic location; after hours accessibility; and physician characteristics (age, sex, country of training).This will be done with the subsets of patients with chronic illness and those with mental illness. Regression analyses (either linear or logistic depending on how the frequency of visits is handled) will be run wherein the unique relationships between models of primary care, geographic location (In accordance with the interests of NSHA, the variable of "management zone" (derived from FSA) has been included in the MAAP-NS data set. This will be used as a measure of geographic location, as will the variables indicating urban/rural categories), accessibility, provider characteristics, and patient status (with regard to mental illness and chronic illness and morbidity index) and primary care utilization are examined.

Methodology for Objective 7

Measures for screening procedures, chronic disease management, and models of primary care were explained above for Objective #4 and #5. For this objective, the rates of screening procedures will be considered outcomes whereas with Objective #5 they will be considered as provider characteristics.

These rates will be compared among the models of primary care; among those with and without afterhours care; and among physician characteristics (age, sex, categories of country of training: Canadian (y/n) or Western (y/n)). The rates will be compared with t-tests or Wilcoxon Rank Sum tests depending on the normality of the data distributions. Three years of data will be used for this objective.

Methodology for Objective 8

Our fax survey to providers asked whether certain specific services were provided: e.g., behavioral counselling, maternity care, mental health care, outreach/home care. These are each represented as yes/no flags. These factors will be compared with patients' use of ED. We will also examine this for a subset of patients with mental illness, and for a subset with selected chronic illnesses.

Service factors will also be compared with patients' use of providers other than their "usual" primary care provider. Again, also looking at this for a subset of patients with mental illness, and for a subset with selected chronic illnesses.

The frequencies will be compared with t-tests or Wilcoxon Rank Sum tests depending on the normality of the data distributions.

Methodology for Objective 9

The practice population estimates obtained from each survey responder will be compared to the number of unique patients linked to each responder in each year of administrative data. Also the average number of patients seen in any two consecutive years and the average number of patients seen over 3 years will be compared to each providers estimate. The analysis will be done using correlation techniques (either Pearson's r for normally distributed data or Spearman's rho for non-normally distributed data).