Reporting of Haemolytic Uraemic Syndrome in the CIHI Hospital Inpatient System and the Alberta Notifiable Disease Registry

Surveillance System Series #1

May, 2005

Health Surveillance, Alberta Health & Wellness

Alberta
HEALTH AND WELLNESS

ACKNOWLEDGEMENTS

EXECUTIVE SUMMARY

Health Surveillance has initiated a series of Surveillance Systems reports to examine data system quality issues. Data systems are the infrastructure that allows surveillance information to be generated through data analysis. Understanding data systems is essential for quality assurance on all analysis, reporting, and interpretation.

This report examines the number of cases of haemolytic uraemic syndrome (HUS) that were reported to the Alberta Communicable Disease Reporting System (CDRS) and to the Canadian Institute for Health Information (CIHI) Hospital Inpatient System – two data systems upon which many provincial and national health statistics are based.

Results show a 3-4 fold difference in the reported cases of HUS between the two data systems in 1994-2003. Findings of this study have implications for the two data systems and for improved reporting of HUS and other diseases from the two systems.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

GLOSSARY

Acronyms

ACCS: Ambulatory Care Classification System

AHCIP: Alberta Health Care Insurance Plan

CDRS: Communicable Disease Reporting System at Alberta Health and Wellness

CIHI: Canadian Institute for Health Information

DOB: Date of birth

E. coli O157:H7: the O157:H7 strain of Escherichia coli bacteria

EHEC: Enterohaemorrhagic Escherichia coli

ESRD: End stage of renal disease

GFR: Glomerular filtration rate

HUS: Haemolytic uraemic syndrome

ICD-9-CM: International Classification of Diseases, Ninth Revision, Clinical Modification

ICD-10-CA/CCI: International Classification of Diseases, 10th Revision, Canadian Modification/the Canadian Classification of Health Intervention

MACAR: Morbidity and Ambulatory Care Abstract Reporting system

STEC: Shiga toxin-producing Escherichia coli

NDR: Notifiable disease report

PHN: Personal health number assigned to an individual's records by Alberta Health & Wellness

QA: Quality assurance

RHA: Regional health authority

TTP: Thrombotic thrombocytopenic purpura

ULI: Unique lifetime identifier

Terms

A confirmed HUS case: A prodrome of enteric symptoms (usually within previous 3 weeks), characterized by acute renal impairment with a higher serum creatinine (>50 μmol/L if <5 years, >60 μmol/L if 5-9 years, >90 μmol/L if 10-13 years, >110 μmol/L if >13 years), microangiopathic haemolytic anemia (Hb<100g/L with fragmented red cells), and thrombocytopenia (<150 000 x 10^9/L) in the absence of septicaemia, malignant hypertension, chronic uremia, collagen or vascular disorders.

A probable HUS case: An acute illness diagnosed as HUS or Thrombotic thrombocytopenic purpura that meets the laboratory criteria in a patient who does not have a clear history of acute or bloody diarrhea in preceding 3 weeks, and has no evidence of invasive streptococcus pneumoniae infection or an acute illness diagnosed as HUS or Thrombotic thrombocytopenic purpura that (1) has onset within 3 weeks after onset of a acute or bloody diarrhea and (2) meets the laboratory criteria except that microangiopathic changes are not confirmed.

A newly reported case of HUS (a proxy of incidence – patients not been tested/yet to be reported are excluded): An individual who was hospitalized for HUS for the first time between January 1, 1994 and December 31, 2003 or who is newly infected with E. coli and reported to CDRS as a HUS case. One patient is counted only once over the 10-year period regardless of the number of hospital admissions.

A prevalent case of HUS: An individual who was admitted into hospital for HUS in a given year between January 1, 1994 and December 31, 2003. One patient may have repeated hospital admissions at different years thus counted more than once over the 10-year study period.

Incidence/ hospital admission ratio: The incident HUS cases over the number of hospital admissions for HUS at a given time period – useful for incidence estimation.

Co-morbidity: Co-existing disease/conditions with HUS that are diagnosed at the same or different time as HUS. Conceptually, co-morbidity is different from complications but this term may be used together with complications in practice.

Dialysis: including hemodialysis, peritoneal dialysis, venous catheterization for renal dialysis.

The first three diagnoses: The diagnoses most responsible for hospitalization, the 1st, 2nd, and 3rd diagnosis was used in the study for consistency and ease of interpretation.

Sensitivity is a measure of the probability of correctly diagnosing/classifying a case or event (i.e., true positive rate).

Specificity is a measure of the probability of correctly diagnosing/classifying a non-case or non-event (i.e., true negative rate).

Positive Predictive Value: In screening and diagnostic tests, the probability that a person/event with a positive test is a true positive.

Overall agreement is a measure of the probability of correctly diagnosing/classifying a case or event plus a non-case or non-event.

Kappa is a measure of the degree of nonrandom agreement between observers or measurements of the same categorical variable. If the measurements agree more than is expected by chance, kappa is positive, otherwise, kappa is negative.

Residence area: Five geographic areas in the province were defined according to reported residence by RHA (2004 boundaries), including Southern Alberta (RHAs 1-2), Central Alberta (RHAs 4-5), Northern Alberta (RHAs 7-9), the Calgary area (RHA 3), and the Edmonton area (RHA 6).

Validity: The extent to which the study measures what it is intended to measure. Practically, validity is the extent to which the results of a method agree with an independent external criterion.

BACKGROUND

This project was initiated by the Health Surveillance Branch, DC&P Branch, and the Provincial Health Office, and particularly by Drs. Stephan Gabos and Karen Grimsrud. The 1st draft was presented across business areas of Alberta Health and Wellness in 2002. This is an update, adding new information for the years 2000-2003 and information about lab-confirmation for HUS. This document is intended primarily for the users of surveillance data systems, pediatricians/other health care providers, hospital administrators, medical health officers, and policy-makers.

INTRODUCTION

The Alberta Communicable Disease Reporting System (CDRS) is a data system for notifiable disease monitoring and assessment. Completeness and accuracy of data is crucial for surveillance activities, therefore quality assurance and improvement efforts are always ongoing. Data from this registry rely on the reporting from regional health authorities. The Canadian Institute for Health Information (CIHI) Hospital Inpatient System is a major national data system for compiling health statistics and monitoring health trends.

A recent analysis by the Health Surveillance Branch (Nguyen, 2000) demonstrated that about 3% of cases with some gastrointestinal infections (giardiasis, campylobactosis, and cryptobacteria) reported by the Canadian Institute for Health Information (CIHI) Hospital Inpatient System was not present in the CDRS.

Questions that arise include: Do differences exist between the CDRS and the CIHI Hospital Inpatient System for other diseases? If so, what is the extent of the difference and what should be done to improve or enhance these databases for surveillance purposes?

As a pilot study, haemolytic uraemic syndrome (HUS) – a rare but severe condition was selected as an additional disease to examine. This report compares cases of HUS reported in the CDRS and in the CIHI Hospital Inpatient System.

CLINICAL AND EPIDEMIOLOGICAL FEATURES OF HUS

Haemolytic uraemic syndrome (HUS) is a clinical syndrome characterized by acute renal failure, microangiopathic haemolytic anemia, and thrombocytopenia (Andreoli et al., 2002; Alberta Health and Wellness, 2003). The average annual reported incidence of HUS in Canadian children younger than 15 years was 1.44 per 100,000, with the peak incidence rate of 3.11 for children younger than 5 years (Rowe et al, 1991). The incidence of HUS in Alberta children was 2.9 times that in Ontario (Rowe et al., 1991).

HUS is the major cause of acute renal failure in infants and young children and is a substantial cause of mortality and chronic morbidity (Siegler, 2003; Cummings, 2002; Andreoli et al., 2002). Clinical manifestations and prognosis of HUS may differ between children and adults (Siegler et al., 1997). About 90 percent of HUS in children is associated with a diarrheal illness (Karmali et al, 1985; Griffin; 1991), 3 to 5 percent of cases may die during the acute phase (Siegler, 1995) and on average death or end-stage of renal disease (ESRD) occurs in about 12% of patients with a maximum follow-up of 22 years (Garg et al, 2003), and 25 to 41 percent of survivors demonstrate long-term mild renal sequelae, such as hypertension, proteinuria, and low glomerular filtration rate (GFR) (Gagnadoux et al., 1996; Garg et al, 2003; Siegler, 2003). HUS in adults may present with heterogeneous clinical manifestations, representing multiple underlying etiologies (Andreoli, 1998; George, 1998). Clinically, patients may present with symptoms ranging from mild diarrhea to hemorrhagic colitis. Kidneys, blood, the brain, liver, heart, lungs, and pancreas may also be involved (Siegler, 1994; George, 1998; Melnyk et al., 1995; Alberta Health & Wellness, 2003). Adult patients, especially elderly individuals, often experience a higher rate of death and disability (Melnyk et al., 1995, Siegler, 2003). Some HUS patients may reoccur after initial recovery (Siegler et al., 1993).

Current evidence suggests that Shiga toxins 1 and 2 are the most important virulence factors for the development of post-diarrheal HUS and are required for disease expression (Siegler, 2003). Shiga toxin-producing Escherichia coli (STEC) infections cause most (85%) postdiarrheal HUS in North American and European children (Goldwater et al., 2000; Karch et al., 1999; Rowe, 1998; Van de Kar, 1996). Among the many STEC serotypes, enterohemorrhagic E.coli O157:H7

2

is the most commonly isolated in North America (Griffin 1991; Rowe, 1998), accounting for at least 80% of STEC infections in the USA (Cummings, 2002). E.coli O157:H7 is a bacterium that exists and multiplies naturally in the intestines of cattle, and perhaps in sheep and other animals (Elder et al., 2000; Pradel et al., 2001). Undercooked ground beef and raw untreated milk are two major food sources of transmission. The organism can also be transmitted through consumption of water (Swerdlow DL et al. 1992; Friedman et al., 1999), vegetables (Ackers et al, 1998) and fruit juice (Cody et al, 1999), or through contact with livestock, infected persons, and contaminated food products (Rowe et al., 1993; Crump et al., 2002).

About 33-50% of diagnosed cases of E. coli O157:H7 infection is admitted to hospital (Bell et al., 1994, Cummings, 2001; Rogers et al., 1986), and usually 2 to 7 percent of those infected develop HUS (Cummings, 2002; Andreoli et al, 2002). The risk increases with decreasing age. In children five years of age or younger, 10 to 14 percent of symptomatic and culture-confirmed E. coli O157:H7 infections progress to HUS (Rowe, 1998; Cummings, 2002). In outbreaks, the proportion of infected E.coli O157:H7 cases that develop HUS may be higher, with rates of up to 14-30% being reported (Bell et al., 1997; Besser et al., 1999; Misselwitz et al., 2003). Although STEC infections and postdiarrheal HUS cases can occur in spatial-temporal circumscribed outbreak settings, most infections and HUS cases occur sporadically (Rowe et al., 1998; Cummings et al., 2002). Between 1998 and 2003, a total of 32 HUS incident cases and 1353 E. coli O157:H7 infections were reported in all ages of Albertans, with an HUS versus E. coli ratio of 0.02 for the total population (Alberta Health and Wellness, 2003, CDRS internal data). In infants however, this ratio is 0.29 - 145 times higher than the total population, suggesting a much closer relationship between E. coli and HUS in children. The major clinical patterns of HUS in Alberta children are microangiopathic hemolytic anemia (Hgb 94 +/- 26 g/L), thrombocytopenia (platelets 87 +/- 83 X 10(9)/L), and acute renal failure (oligoanuria with a BUN of 26 +/- 15 mmol/L, and a creatinine of 294 +/- 90 mumol/L). The key manifestations in Alberta children were: diarrhea (100%), vomiting (80%), hemorrhagic colitis (79%), abdominal discomfort/tenderness (59%), elevated hepatocellular enzymes (58%), indirect hyperbilirubinemia (49%), fever (33%), rectal prolapse (13%), colonic stricture (3%), colonic perforation (1%), and intussusception (1%). Elevation of amylase and lipase in the presence of acute renal failure is also presented (Rowe, 1991).

There is no known therapy to halt the progression of HUS. The active stage of the disease usually lasts one to two weeks, during which a variety of complications are possible. With possibility of both acute renal failure and long-term complications, HUS imposes a significant burden on the health care system and families involved. For instance, in Alberta 56.6% of children with HUS required dialysis, with 17 days of hospitalization, and about 5.3% of cases died of complications attributable to HUS (Grodinsky et al., 1990; Rowe 1991). In the United States, an HUS patient stay in hospital as long as 325 days, with a median length of stay of 11 days, and a median charge for hospital care of $39,508 USD per child (Cummings, 2002). About 3-5% may develop ESRD within the first few years (Siegler, 2002) and thus require chronic renal replacement therapy (dialysis or renal transplantation) which is very costly. About 30-50% of patients develop less severe renal sequelae, and about 10% of this group may eventually develop ESRD due to hyperfiltration injury (Siegler, 2002). In some countries, about 50% of patients require dialysis due to kidney failure, 25% develop pancreatitis, 25% experience seizures, and 5% suffer from diabetes mellitus (http://www.about-hus.com, 2003). The majority of HUS patients require transfusion of blood products and develop complications common to the critically ill. The prognosis of HUS is improving with better identification and treatment (Garg et al., 2003).

STUDY QUESTIONS

The current analysis was guided by the following questions:

1. What are the sources of diagnostic information for HUS in the CIHI Hospital Inpatient System and in the CDRS? Is the diagnosis of HUS reliable in the CIHI Inpatient System?
2. Is there a difference in reported cases of HUS between the CDRS and the CIHI Inpatient System?
3. What is the regional distribution of reported HUS in Alberta?
4. Are there differences between hospitals in the reporting of HUS?

MATERIALS AND STUDY METHODS

Data Sources

The following data were used:

❑ *Canadian Institute for Health Information (CIHI) Hospital Inpatient System*. This system contains inpatient data from all acute care hospitals in Alberta, about 350,000 hospital separations each year. Data are routinely collected by hospitals and reported to regional and provincial health authorities and to CIHI for funding and monitoring purposes. Information available may include patient demographics, residence, chart number, diagnoses, date(s) of admission and separation, service provider(s), service(s) provided, service delivery site(s) and facility type, resources, and the like. The data collected in the system are increasing over time. Currently, the system contains a total of 25 diagnoses responsible for hospitalization compared to 16 during 1992/93 and 2001/02, and to 3 prior 1992/93. Since 2001, the CIHI Hospital Inpatient System in Alberta become part of the MACAR application which uses both inpatient and outpatient data. This study used inpatient data with discharges between January 1, 1994 and December 31, 2003.

❑ *Communicable Disease Reporting System* (CDRS), January 1994 – December 2003. Communicable disease case reports are centrally collected and maintained in the CDRS, a secure database managed by the Disease Control and Prevention Branch of Alberta Health and Wellness. Positively confirmed laboratory reports are required to complete each notifiable disease record. Laboratory reports are submitted directly from the Provincial Laboratory for Public Health or the Medical Officer of Health to the DC&P Branch. A case in principle only becomes counted in the CDRS upon receipt of both a confirmatory laboratory report and a Notifiable Disease Report (NDR).

❑ *Alberta Health Care Insurance Plan (AHCIP) Stakeholder Registry*. The Alberta Health Care Insurance Plan Stakeholder Registry was established to enable premium collection and assessment of registrant eligibility for services claimed by medical practitioners. This registry covers virtually all residents of the province except a small proportion of special population groups (i.e. members of the Armed Forces and RCMP, federal inmates, persons from other provinces during their first three months in Alberta). The registry collects and maintains demographic, socioeconomic, residential, migration, and other

information. Several data files are derived from this registry, including mid-year population, fiscal year-end cumulative population, and cumulative stakeholder registration. This registry is linkable, at the individual level, to all health care databases maintained by Alberta Health and Wellness. The current report used demographic information, current as of Feb 2004.

Diagnostic and health care service procedure information in the CIHI Hospital Inpatient System in Alberta prior to April 2002 was coded by the *International Classification of Diseases, Ninth Revision, Clinical Modification* (ICD-9-CM). Since April 2002, this information was coded by *Classification of Diseases, tenth Revision, Canadian Modification (ICD-10-CA) and the Canadian Classification of Health Intervention (ICD-10-CA/CCI). Diagnosis in the CDRS was coded by ICD-9-CM.*

Case Identification, Definition, and Assessment

HUS falls into the diagnostic category of non-autoimmune hemolytic anemia. To be able to distinguish HUS from other conditions of this category, the fifth digit of the ICD-9-CM (ICD-9-CM=283.11) is required. The CIHI Inpatient System has detailed ICD-9-CM codes up to the fifth digit, complete since 1994. In ICD-10-CA, a 4-digit code (ICD-10=D59.3) is required to identify HUS. In CDRS, HUS was reportable since 1983, based on a positive lab test of E. coli O157:H7 and/or an NDR. The case definition for HUS was developed and implemented recently (Alberta Health and Wellness, 2003).

Case Definition: An HUS case was defined as a patient who was hospitalized at least once with a diagnostic code of HUS (ICD-9-CM=283.11 or ICD-10-CA=D59.3) in a year. Since HUS can lead to or may result from several complications, multiple sources of diagnosis in case ascertainment are required. The first three diagnoses responsible for hospitalization were used to extract the diagnostic information of HUS for the ease of data analysis and consistency during the study period. Cummings et al (2002) suggested that the majority of HUS cases are captured from the 1st three diagnoses. Since most cases with HUS are hospitalized (Rogers 1986), population-based HUS incidence derived from hospitalization data is expected to be a close approximation to actual rates (Rowe et al, 1991).

To assist case assessment, health care service information was also extracted for the period 1994 and 1999. The first three procedure codes (ICD-9-CM) were used in this assessment. About 44 percent of records did not have a service code. For cases with a service code, the major categories of the services provided are summarized in *Table 1*.

Table 1 Percent distribution of service procedures for HUS, CIHI data, 1994-1999

Description of Procedure	ICD-9 Procedure Code	%
Creation of cutaneoperitoneal fistula	5493	17
Transfusion of blood & blood components	9904, 9907, 9915	13
Venous catherization	3893	16
Renal dialysis	3895, 3927, 3942	10
Hemodialysis	3995	8
Kidney Transplant	5569, 5553	8
Peritoneal dialysis	5495, 5498	3
Renal biopsy	5523, 5524	2
Insertion of vascular access device	8607	2
Other procedures	0331, 3324, 3451, 3491, 3891, 3927, 3942, 4101, 415, 4516, 4524-4525, 4573, 5419, 544, 741, 7569, 8605, 8703, 8809, 8819, 8872, 8952, 9396, 966, 9671-9672, 9921, 9971	21

As shown, most of these services are likely to have a direct or indirect relation to the treatment of HUS and its complications. It is assumed that if a patient was diagnosed with HUS, then an HUS-related treatment would have been provided.

Measure of the Completeness of Data and Validity of Diagnosis

The completeness of data is defined here as the extent of data collection and/or reporting in a surveillance system regardless of the precision or accuracy of the data. The validity of a diagnosis is the degree to which a procedure or protocol is able to measure what it is intended to measure. A more pragmatic definition, used here, is 'the extent to which the results of a method agree with an independent external criterion' (Bennett et al., 1975).

Sorensen and his colleagues (1996) suggested the following approaches for the evaluation of the completeness and reliability of data from administrative data sources:

- Comparison of the diagnosis to an accepted standard ('gold standard') or to one or more independent reference sources
- Comparison of the total number of cases across data sources
- Comprehensive review of health records/patient charts
- Interview/survey of the cases diagnosed
- Re-examination of the cases diagnosed

This study compared the hospital diagnosis to the CDRS diagnosis at both individual and aggregated levels. The completeness of registry of HUS in the CDRS is evaluated by looking at: the difference in the total number of cases (the CIHI minus the CDRS), and the case ratio (the CIHI divided by the CDRS).

The interpretation is:

- difference = 0 or ratio = 1:   Equal-reporting in the CDRS
- difference > 0 or ratio > 1:   Under-reporting in the CDRS
- difference < 0 or ratio < 1:   Over-reporting in the CDRS

The agreement between the CDRS and the CIHI Hospital Inpatient System for diagnoses of HUS was also examined. The reported HUS cases in the CDRS are based on lab report and/or clinical evaluation, thus the diagnosis from the CDRS was used as the "gold standard" for validity comparisons. The overall percent agreement, sensitivity, positive predicative value (PPV), and kappa were estimated according to Fleiss (1981), Altman (1995), and Sorensen et al (1996). The definition and calculation of each measure are illustrated **Appendix A**. The criteria used to judge the level of agreement are: greater than 80% for excellent, 61-80% for good, 41-60% for moderate, and less than 40% for poor (Wang et al., 1994).

Measure of Co-Morbidity

A HUS patient may have other disease(s)/conditions thus be given other diagnoses at the time of admission or may progress to other conditions/diseases over time. For each new HUS case identified from the CIHI Hospital Morbidity Inpatient System between 1994 and 2003, other diagnoses at admission and all subsequent hospital admissions were extracted for follow-up until March 31, 2004 by the year of first diagnosis of HUS. The first three diagnoses for hospital admission were extracted and grouped by selected disease grouping and ICD-10 chapter.

Data Linkage and Analysis

To compare the diagnosis from different data sources for each individual, the cases across data sources must be matched by a predefined linkage criteria and linkage protocol. The unique lifetime identifier (ULI) was used for performing the linkage. Between 1994 and 2003, a total of 58 new HUS cases were reported to the CDRS, 42 (72.4%) cases did not have a ULI. Using the names, birth date and sex, 40 of these 42 cases were successfully assigned a ULI according to the AHCIP Stakeholder Registry. Of the 56 cases with a ULI, 51 (91.1%) were matched to the HUS database extracted from CIHI Inpatient System by ULI. As a quality assurance (QA) measure, patient's date of birth (DOB) and sex were used to verify the accuracy of the linkage. All 56 patients with ULI had identical DOB and sex between the CIHI Hospital Inpatient System and the CDRS. The two cases from the CDRS without a PHN were excluded from the validity analysis.

Descriptive analysis was performed. Because of the small number of cases, the regional data is presented for five areas of the province: Edmonton area, Calgary area, Northern Alberta, Central Alberta, and Southern Alberta. A Chi-square test was applied as appropriate for regional variations and for trend analysis (Ataman, 1995).

MAJOR FINDINGS

Newly Reported Cases from CDRS

Most (91.2%) of the HUS cases reported to the CDRS are 0-17 years, 10.5% of them are infants, and 68.4% are children four years or younger. Overall, 62.1% of cases had reported a lab confirmation (of E. coli infection) in 1994-2003, fluctuating from 44.4% in 2004 to 100.0% in 1999 (**Appendix B**). Fifty-three (91.4%) cases were reported as hospitalized, four unknown, one not hospitalized.

Newly Reported Cases from CIHI

Figure 1 *Distribution of sources of diagnostic information of HUS, CIHI data, 1994-2003*

Overall, a total of 238 HUS incident cases and 391 hospitalizations for HUS were identified between January 1994 and December 2003 in Alberta, with a case/hospitalization ratio of 0.61. Most (74%) of the reported cases had only one hospital admission per year, with a maximum of four hospitalizations each year. The median of the length of stay was 6-days, ranging from a few hours

to a maximum of 179 days. Overall, the majority (65.0%) of the cases were captured from the most responsible (1st) diagnosis (**Figure 1**). However, this distribution varies by age group. For children under 15 years, the HUS cases captured from the 1st diagnosis is 79.1%, while the corresponding proportion for cases in ages 15-64 years and 65 year or over was only 37.9% and 23.5%, respectively (**Figure 2**). For these two age groups, kidney failure, hypertensive renal disease, and fluid volume depletion – common conditions and complications of HUS are primary reasons for admission.

Figure 2 *Source of diagnostic information of HUS by age group, CIHI data, 1994-2003*
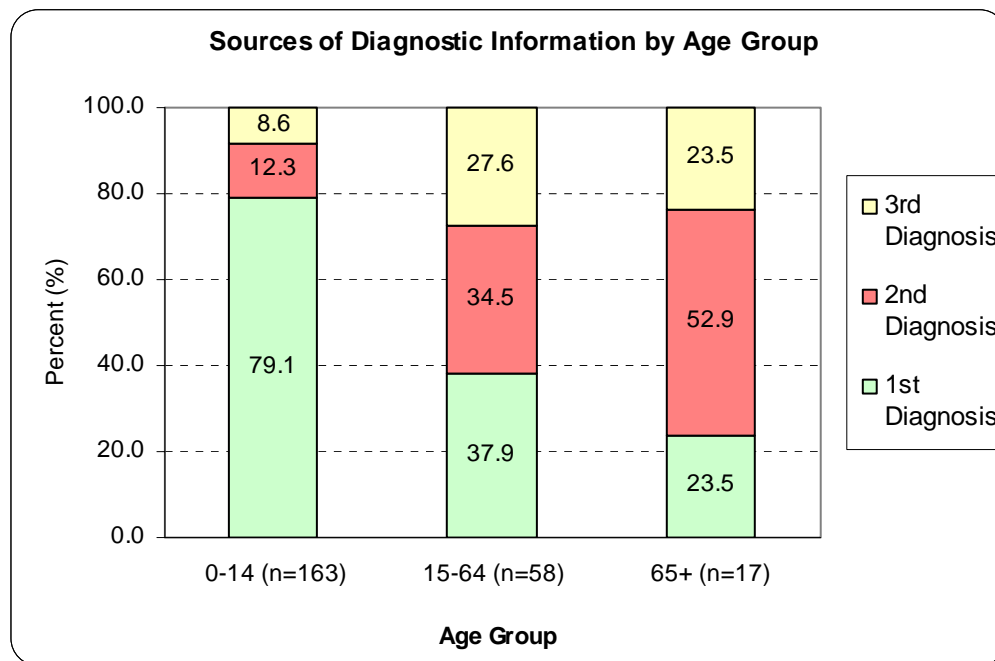
Figure 3 shows the percent distribution of sources of diagnostic information of HUS for cases matched and not-matched between the CDRS and the CIHI Hospital Inpatient System. A total of 51 cases were matched between the two data systems. Compared with data from the CDRS, about 91% of cases reported in the CDRS were captured by the CIHI Hospital Inpatient System.

Figure 3 *Source of diagnostic information of HUS by study group, CDRS vs. CIHI, 1994-2003*

Surveillance System Series – Reporting of HUS in CIHI and CDRS
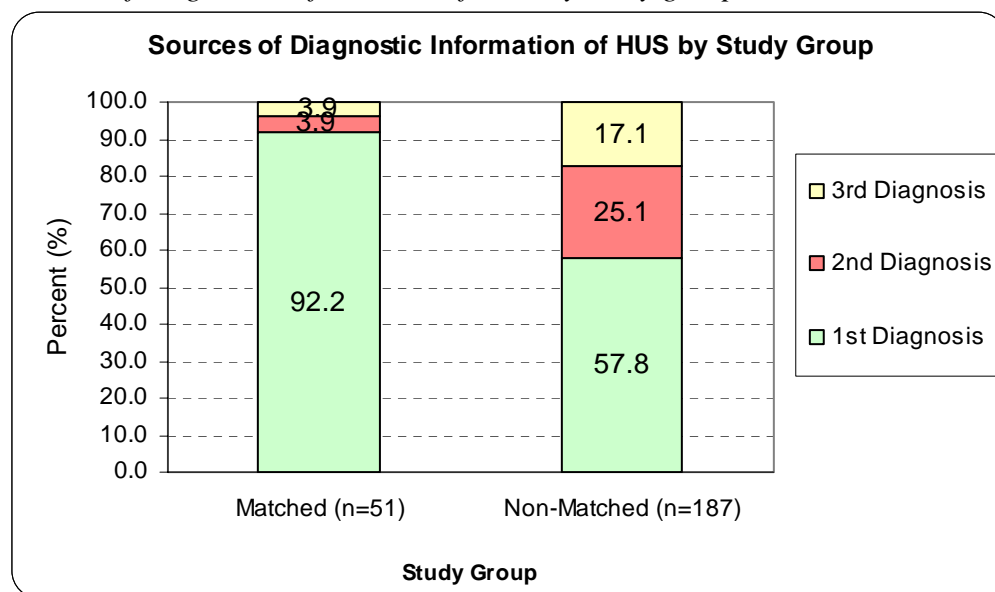
For these captured cases, 92.2% were from the 1st diagnosis, 3.9% from the 2nd and 3rd diagnosis, respectively. In contrast, for those reported by the CIHI Hospital Inpatient System but not by the CDRS, the percentage was 57.8%, 25.1%, and 17.1% for the 1st, 2nd, and 3rd diagnosis, respectively. These findings suggest about 42% of HUS cases are not reported to the CDRS likely due to case ascertainment from only the 1st diagnosis in practice.

Of the 51 matched cases from the CDRS, 16 (31.3%) had repeated hospitalizations and seven (13.7%) were reported to the CDRS not at their first time of hospital admission but for subsequent ones, and one died two days after first admission.

Sex and Age Distribution of Newly Reported HUS Cases

Of 238 incident cases reported from the CIHI Hospital Inpatient System, 143 were females and 95 were males, with a female/male ratio of 1.5. The high female/male ratio is primarily driven by the cases aged between 15 and 64 years (**Table 2**). About 4.0% of HUS cases from the CIHI were infants less than one year old – less than half as that from the CDRS, 63.0% were children less than 10 years, and 71% were children 17 years or younger.

Table 2 *Age distribution of newly reported HUS by sex in Alberta, CIHI, 1994-2003*

Age Group	Female		Male		Female/Male
	N	%	N	%	Ratio
0-14	85	59.4	65	68.4	1.3
15-64	48	33.6	22	23.2	2.2
65+	10	7.0	8	8.4	1.3
All Ages	143	100.0	95	100.0	1.5

The age distribution varies by the source of diagnosis (**Figure 4**). The proportion of cases in children under 15 years was higher (83.2%) for cases identified from the first diagnosis, but lower for those from the second (40.8%) and third (41.2%) diagnosis. This finding suggests about 60% of HUS cases identified from the second and third diagnosis are patients 15 years or older. If only the primary diagnosis is used, these patients will be excluded from reporting. The detailed breakdown of HUS cases by sex, age group and sources of diagnosis from the CIHI data is presented in **Appendix C**.

Figure 4 *Age distribution by source of diagnosis, CIHI 1994-2003*

Age Distribution by Source of Diagnosis
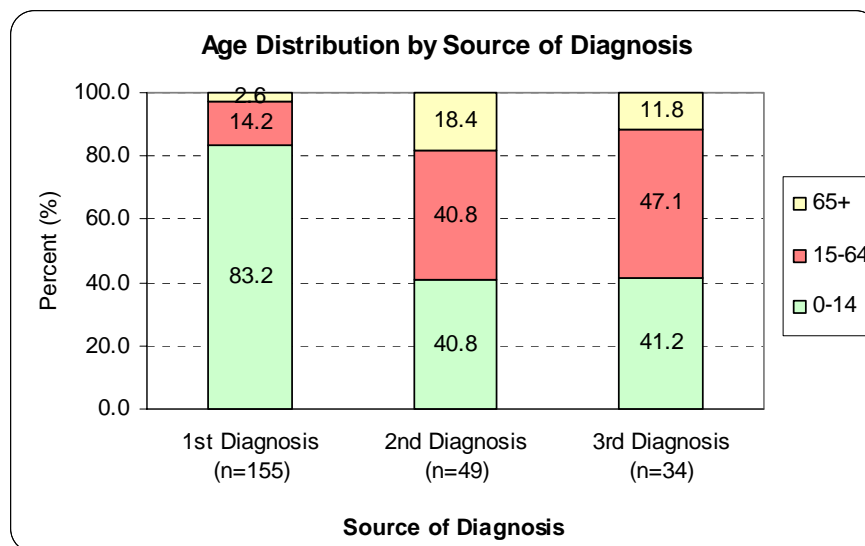
Completeness of HUS Registration in CDRS, 1994-2003

Table 3 shows a year by year comparison of HUS cases in all ages reported in the CIHI Hospital Inpatient System and the CDRS. Compared to the data from the CIHI, there are, at least, 180 (75.6%) HUS cases that may be under-reported in the CDRS during 1994-2003. During this period, the number of cases from the CIHI is on average 4.1 times of that from the CDRS.

When the comparison is limited to only children 0-17 years, the pattern of underreporting from the CDRS is still evident but to a lesser degree with 115 (68.9%) cases underreporting (**Table 4**). This finding agrees with a recent study in the United States, which reported a 56.1% of cases underreporting (Cummings, 2002).

Table 3 *Number of newly reported cases with HUS in all ages between the CDRS and CIHI Hospital Inpatient System, Alberta, 1994-2003*

Year	Cases from CDRS [a]	Cases from CIHI [b]	Differences: CIHI vs. CDRS	
			CIHI minus CDRS	CIHI/CDRS Ratio
1994	9	28	19	3.1
1995	5	17	12	3.4
1996	6	25	19	4.2
1997	6	27	21	4.5
1998	2	16	14	8.0
1999	1	25	24	25.0
2000	6	18	12	3.0
2001	5	30	25	6.0
2002	9	26	17	2.9
2003	9	26	17	2.9
1994-2003	58	238	180	4.1

[a] *Lab confirmed and/or with a NDR report as haemolytic uraemic syndrome (ICD-9-CM = 283.11, ICD-10=D59.3).*

[b] *Extracted from the first three diagnoses (ICD-9-CM = 283.11, ICD-10=D59.3) of the CIHI system.*

Table 4 *Number of newly reported cases with HUS in children 0-17 years between the CDRS and CIHI Hospital Inpatient System, Alberta, 1994-2003*

Year	Cases from CDRS [a]	Cases from CIHI [b]	Differences: CIHI vs. CDRS	
			CIHI minus CDRS	CIHI/CDRS Ratio
1994	9	17	8	1.9
1995	5	9	4	1.8
1996	6	21	15	3.5
1997	6	15	9	2.5
1998	2	12	10	6.0
1999	1	18	17	18.0
2000	6	15	9	2.5
2001	5	24	19	4.8
2002	5	19	14	3.8
2003	8	17	9	2.1
1994-2003	53	167	114	3.2

[a] *Lab confirmed and/or with a NDR report as haemolytic uraemic syndrome (ICD-9-CM = 283.11, ICD-10=D59.3).*

[b] *Extracted from the first three diagnoses (ICD-9-CM = 283.11, ICD-10=D59.3) of the CIHI system.*

It should also be noted that the cases reported by the CDRS are all confirmed with a Notifiable Disease Case Report and often (65% in 2000-2002) with a lab-confirmation of E. coli O157:H7 infection. However, not every case is necessarily reported. In practice, only the primary (1st) diagnosis is used as the source of case ascertainment. In the present study, the primary diagnosis accounted for only 65% of incident HUS cases reported in the CIHI Hospital Inpatient System.

Accuracy of the Information - Validity of Diagnosis

Table 5 presents the overall agreement, sensitivity, positive predictive value, and kappa on reported diagnosis of HUS by the CIHI Hospital Inpatient System. As shown, the sensitivity is reasonably high (91.1%) but the overall agreement is poor (21.0%). Although about 91% of HUS cases from the CDRS would be picked up by the CIHI Hospital Inpatient System, only about 21% of HUS cases classified by the CIHI may be supported by the CDRS. The negative kappa value suggests this agreement could be due to chance alone. These findings indicate a large disagreement in newly reported cases of HUS between the CIHI and the CDRS between 1994 and 2003.

Table 5 *Validity of diagnosis of HUS among cases of all ages in the CIHI Hospital Inpatient System, Alberta, 1994-2003*

Cases from CIHI	Cases from CDRS		Measure of Internal Validity			
	Yes	No	Overall Agreement	Sensitivity	PPV	Kappa
Yes	51	187				
No	5	0	21.0	91.1	21.4	-0.1
Total	**56**	**187**				

PPV - Postive predictive value.

Severity of Illness

Among the 133 cases with data on urgency of admission indicator in the CIHI Hospital Inpatient System, 117 (88.0%) required immediate attention. Of 238 incident cases, 8 (3.4%) cases were reported as dying in the hospital at the time of diagnosis. This hospital case fatality was higher than that (2.7%) reported by Cumming and his colleagues (2002). When we followed-up the case cohort of each year in their subsequent hospital admissions, with a range of one to ten year follow-up, a total of 23 of these HUS cases died in hospitals. A subsequent linkage analysis shows 60.0% of the HUS hospital deaths occur within the same year since the 1st hospitalization, 16.0% the second year, and 24.0% in 3-10 years. The majority of HUS cases may live for a minimum of 10 years or longer after first hospital admission.

*Co-morbidity: Associated Diagnoses and Complications*

Co-morbidity was estimated by comparing the diagnostic categories from the three sources of diagnoses in the CIHI Hospital Inpatient System between 1994 and 2003. During the 10-year study period, of 238 new HUS cases there were 346 subsequent hospital admissions for HUS, including 182 for kidney failure, 55 for fluid volume depletion, 69 for E. coli infection, and 12 for diabetes. When using only primary diagnosis for hospital admissions, there were 66 admissions for respiratory disorders, 52 for circulatory system disorders, 49 for injury and poisoning, 47 for digestive system disorders, 46 for genitourinary system disorders, 36 for symptoms and signs, 32 for infectious disease, 27 for mental disorders, and 20 for endocrine and metabolic disorders between 1994 and 2003. While we can not assume these subsequent hospital admissions after HUS diagnosis result from HUS, a wide range of co-morbidity with HUS is evident.

For those coded with E. coli infection in the CIHI Hospital Inpatient System, the proportion of source of diagnostic information was 59.8%, 31.6%, and 8.6% from the first, second, and third diagnosis responsible for hospitalization in 1994-2003, respectively. Of a total of 174 hospital admissions for E. coli in 1994-2003, 69 (39.7%) had a diagnostic code of HUS, and 13 (18.8% of HUS) are from 2nd and 3rd diagnoses. In 2003/04, there are a total 35 hospitalizations for E. coli from first diagnosis. Of these, four admissions also had a diagnosis of HUS from the second, third, fourth, and seventh diagnosis, respectively. If only the first diagnosis is used, these four HUS cases associated with E. coli infection would not be captured/reported. In the current study, 83 (34.9%) cases of HUS would have been missed if only the primary diagnosis was used for case identification and definition.

Regional Distribution of Newly Reported HUS Cases

The regional distribution of HUS is presented for five areas in the province (**Figure 5**). For cases captured by both the CDRS and CIHI Hospital Inpatient System (matched), about 73% were residents of Calgary region or of southern Alberta. In contrast, for cases reported only by the CIHI Hospital Inpatient System (non-matched), residents of Edmonton region had the largest proportion (34.3%), followed by Calgary region (26.3%), northern Alberta (16.0%), Central Alberta (13.1%), and southern Alberta (10.3%). The detailed data by regional health authority is attached in **Appendix D**.

Figure 5 *Newly reported HUS cases by area of residence and study group in Alberta, CIHI, 1994-2003*
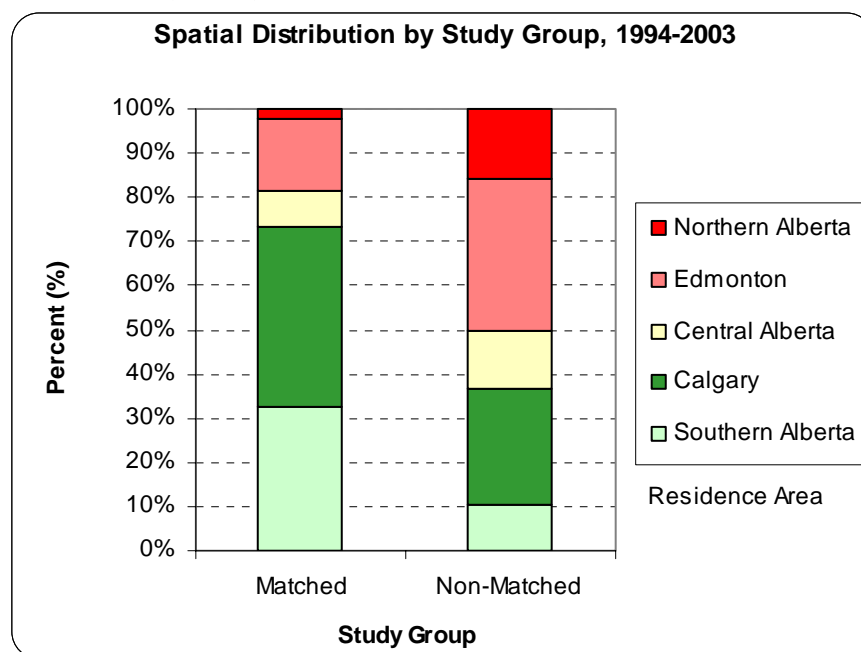
For cases living in Calgary region or southern Alberta, about 30% and 47% were respectively reported or captured by the CDRS (the matched group), higher than the 21% for the provincial average (**Table 6**). In contrast, for cases living in Edmonton region, northern or central Alberta, only about 3%-15% of them were reported/captured by the CDRS.

Table 6 *Distribution of newly reported HUS cases by residence area in matched and non-matched group, Alberta, 1994-2003*

Residence Area	Matched Group		Non-Matched Group		Total Cases
	N	%	N	%	
Southern (RHAs 1, 2)	16	47.1	18	52.9	34
Calgary (RHA3)	20	30.3	46	69.7	66
Central (RHAs 4,5)	4	14.8	23	85.2	27
Edmonton (RHA 6)	8	11.8	60	88.2	68
Northern (RHAs 7,8,9)	1	3.4	28	96.6	29
Unknown RHA	1	7.1	13	92.9	14
Total	**50**	**21.0**	**188**	**79.0**	**238**

Distribution of Newly Reported HUS Cases by Hospital Facility, 1994-1999

This section uses data from previous analysis, because higher reporting by hospitals in Calgary and southern Alberta remains the same, and changes in the institution number and in the number of facilities in the Province make it difficult to combine the data over time.

Of 117 HUS cases between 1994 and 1999, 65 (55.5%) were reported by hospitals in Calgary, southern and central Alberta, and 52 (44.5%) reported by hospitals in Edmonton region and northern Alberta (**Appendix E**). All cases but one captured by the CDRS were reported by hospitals in southern and central Alberta. The proportion of cases who were reported by both the CDRS and the CIHI Inpatient System (the matched group) was generally higher for hospitals of southern and central Alberta, with an average of 38.5%. The proportion was particularly higher for Lethbridge Regional Hospital (75%) and Alberta Children's Hospital (55.6%). In contrast, the proportion in hospitals of the Edmonton area and northern Alberta was much lower. The Royal Alexandra hospital reported one case (20%) that was present in both the CDRS and the CIHI Hospital Inpatient System. The rest of cases who were reported by hospitals in Edmonton area and northern Alberta were all not reported to the CDRS.

Findings of general lower percentage in the matched group by hospitals in the Edmonton area and northern Alberta are consistent with those from **Figure 5** and **Table 6**. If the cases identified through the CIHI Inpatient System are misdiagnosed or not confirmed then they may not be reported to the CDRS by hospitals. Only two residents in the Edmonton area and northern Alberta were in the matched group (Table 6), one of whom was reported by hospitals in Edmonton. On the other hand, if the cases identified through the CIHI Hospital Inpatient System are true HUS patients then the observed lower percentage in the matched group may suggest a significant under reporting of HUS to the CDRS by hospitals in the Edmonton region and northern Alberta.

DISCUSSION

This pilot study, the first step of evaluation and development of the CDRS initiative, compared data on reported HUS cases between the CDRS and the CIHI Inpatient System. It was found that 91.17% of the cases recorded in the CDRS were captured by the CIHI Hospital Inpatient System. However, the overall agreement between the CDRS and the CIHI Inpatient System was poor (21.0%) and there were significant differences in reported HUS cases between the two data sources. Compared to the CIHI Inpatient system, the CDRS had an average under-reporting in 1994-2003, at least, of 75% for all ages, and of 69% for children under 18 years. This under-reporting is higher than that reported in the United States (56.1%).

Why were there such large differences in reported HUS between the CDRS and the CIHI Inpatient System? While there may be many possibilities, the validity of diagnosis and completeness of reporting are the two areas that warrant attention.

Validity of Diagnosis

A case reported by the CDRS is usually lab-confirmed and/or with a NDR case report. However, a case identified through the CIHI Inpatient System in the present study was based on the diagnostic codes available in the system. Although this approach has been used by others (Cummings et al., 2002; Rowe et al., 1991), the validity of diagnosis of these cases was not adequately assessed in this pilot study. Although HUS is a serious condition and most HUS cases are likely to be admitted into hospital (for treatment) where diagnosis is often confirmed by lab tests, there are still several problems in using discharge diagnoses (Steinberg et al., 1990; Yao et al., 1999), including

1. variations in coding,
2. errors in coding,
3. incompleteness in coding, especially of co-morbidities,
4. limits in the specificity of available codes, and
5. errors and variation in diagnosis (cross physicians and hospitals).

There is no evidence that the data from the CIHI Hospital Inpatient System can be an exception. Inconsistencies in the number of diagnoses and in the ICD coding over time are among such example. Thus, the validity of the diagnosis of HUS reported by the CIHI system should be adequately evaluated and must be considered in the interpretation of the observed differences in HUS cases between the two data systems. It should be noted that the incidence/hospitalization ratio (0.61) observed in the present study is fairly close to that (0.63) reported by Cummings and colleagues (2002), implying the stability of this ratio, thus it may be used in the estimation of HUS incidence.

Completeness of Reporting

Completeness of reporting of HUS should be considered for both the CDRS and the CIHI Hospital Inpatient System. Not all cases with HUS may be reported to the CDRS. Cummings and his colleagues (2002) found only 43.9% of HUS cases in children identified through hospital discharge data system were officially reported to public health authorities, this proportion is higher than 31.7% observed in the current study. In Alberta, the CDRS very likely capture cases identified only by the primary diagnosis. Such a process will naturally lead to an underreporting. For instance, if a patient's primary diagnosis is E. coli O157:H7 infection but the secondary diagnosis is HUS, these potential HUS cases may not have a chance to be further verified by lab-tests, are thus not captured by the CDRS. For the CIHI Hospital Inpatient System, a patient with HUS may not be diagnosed and/or coded as HUS due to errors in diagnosis, coding and data entry. Also, only the first three diagnoses were used for case identification and definition in the present study - likely leading to an underreporting. As reported by others (Cummings et al., 2002), a total of 94.7% of HUS hospital discharge in children under 18 years were captured in the first five diagnoses (among 25 diagnoses available). As such, using only the first three diagnoses in the current study may have resulted in an underestimation of HUS incidence by, at least, five percent.

In summary, the following possibilities should be considered in interpretation of the observed differences between the CDRS and the CIHI Hospital Inpatient System:

- Errors and variations in clinical diagnosis by both data sources

- Incompleteness of coding and reporting of HUS cases in the CIHI Inpatient System or in the CDRS

- Limits in the specificity of available ICD-9 and ICD-10 codes

- Variations and errors in coding/assigning (ICD-9 or ICD-10) diagnoses between coders and physicians and within the coder group for both the CIHI Inpatient System and the CDRS

- Error in data entry and updating

- Differences in record keeping (i.e., incident case or prevalent case, person-based or event-based, etc.)

- Co-morbidities

- Patient migration and loss to follow-up

RECOMMENDATIONS

The general recommendations for the operation of the CDRS and CIHI Hospital Inpatient Systems are to:

1. Enhance the quality assurance on the process of data collection, verification, updating, and reporting.
   a) Establish a regular process for the review of case ascertainment and reporting to the CDRS.
   b) Establish processes to ensure on-going quality assurance for the CDRS and the CIHI Hospital Inpatient System.
   c) Periodically review the process of case collection and reporting to the CDRS and CIHI.
2. Increase the utilization of the CDRS and CIHI Hospital Inpatient System to its full potential value for surveillance and research.

Regarding the specific findings surrounding HUS it is recommended that further investigation into the patterns reported here be initiated including:

1. Conduct a comprehensive record review: a 20% random sample of the HUS cases may be selected from the CIHI Hospital Inpatient System; the hospital charts of the sample may be reviewed.

2. Conduct a survey of all hospitals in the province about the admission and coding practices for HUS and related conditions.

3. Follow-up the HUS cases at 5, 10, and 15 years after diagnosis to analyze the patterns of health care utilization, cost, and treatment and survival of these patients.

Regarding the general reporting of incidence and prevalence rates of HUS, the following is recommended:

1. Use administrative data as a complementary to the CDRS for case identification, case definition, and incidence estimation, particularly for diseases of less severity. More specifically:

    a) Use CIHI Hospital Inpatient System as the major data source of administrative data, and when possible using ACCS and Claims data as supplementary for case identification and case definition.

    b) Use all three diagnoses in case identification and definition, although the primary diagnosis appeared to be more reliable.

    c) Develop a set of case definitions by the level of likelihood, such as a probable case, likely case, and possible case. An incident and prevalent case must be clearly defined.

    d) Develop a standard protocol including computer programs for the calculation of the incidence, prevalence, and case fatality rate.

2. Establish a HUS registry by using both administrative and CDRS data and take an active approach for HUS surveillance, including the evaluation of outcomes of the current medical interventions

3. Education of health professionals reporting of HUS as a notifiable disease independent of E. coli. O157:H7.

REFERENCES

1. Ackers ML. Mahon BE. Leahy E. Goode B. Damrow T. Hayes PS. Bibb WF. Rice DH. Barrett TJ. Hutwagner L. Griffin PM. Slutsker L (1998). An outbreak of Escherichia coli O157:H7 infections associated with leaf lettuce consumption. Journal of Infectious Diseases; 177(6):1588-93.

2. Alberta Health & Wellness (2003). Alberta Case Definitions Manual. Edmonton, AB: Alberta Health & Wellness. June 2003. ISBN: 0-7785-2405-1.

3. Alberta Health & Wellness (2004). CDRS report. As of November 30, 2004. Alberta Health & Wellness Internal Statistics.

4. Andreoli SP. Trachtman H. Acheson DW. Siegler RL. Obrig TG (2002). Hemolytic uremic syndrome: epidemiology, pathophysiology, and therapy. *Pediatric Nephrology*; 17(4):293-8.

5. Altman DG. 2 x K tables - Comparison of several proportions. In DG Altman ed. Practical Statistics for Medical Research. London/Glasgow/Weinheim/New York/Tokyo/Melbourne/ Madras: Chapman & Hall, 1995.

6. Bennett AE, Ritchie K. Questionnaires in Medicine. p. 29. Oxford University Press: London (for the Nuffield Provincial Hospitals Trust) 1975.

7. Besser RE. Griffin PM. Slutsker L (1999). Escherichia coli O157:H7 gastroenteritis and the hemolytic uremic syndrome: an emerging infectious disease. [Review, Tutorial] *Annual Review of Medicine*; 50:355-67.

8. Cody SH, Glynn MK, Farrar JA, Cairns KL, Griffin PM. Kobayashi J. Fyfe M. Hoffman R. King AS. Lewis JH. Swaminathan B. Bryant RG. Vugia DJ (1999). An outbreak of Escherichia coli O157:H7 infection from unpasteurized commercial apple juice. Ann Intern Med; 130:202-09.

9. Crump JA. Sulka AC. Langer AJ. Schaben C. Crielly AS. Gage R. Baysinger M. Moll M. Withers G (2002). Toney DM. Hunter SB. Hoekstra RM. Wong SK. Griffin PM. Van Gilder TJ. An outbreak of Escherichia coli O157:H7 infections among visitors to a dairy farm. [see comments.]. NEJM; 347(8):555-60.

10. Cummings KC. Mohle-Boetani JC. Werner SB. Vugia DJ (2002). Population-based trends in pediatric hemolytic uremic syndrome in California, 1994-1999: substantial underreporting and public health implications. *Am J Epidemiol*;155(10):941-8.

11. Elder RO, Keen JE, et al (2000). Correlation of Enterohemorrhagic Escherichia coli O157 prevalence in feces, hides, and carcasses of beef battle during processing. Proc Natl Acad Sci USA;97:2999-3003.

12. Fleiss JL (1981). Statistical Methods for Rates and Proportions 2nd ed. New York: Wiley; pp. 44-46, 212-225.

13. Friedman MS, Roels T, Koehler JE, Feldman L, Bibb WF, and Blake P (1999). Escherichia coli O157:H7 Outbreak Associated with an Improperly Chlorinated Swimming Pool. Clinical Infectious Diseases;29:298-303.

14. Gagnadoux MF. Habib R. Gubler MC. Bacri JL. Broyer M (1996). Long-term (15-25 years) outcome of childhood hemolytic-uremic syndrome. Clinical Nephrology. 46(1):39-41.

15. Garg AX, Suri RS, Barrowman N, Rehman F, Matsell D, Rosas-Arellano MP, Salvadori M, Haynes RB, Clark WF (2003). Long-term renal prognosis of diarrhea-associated hemolytic uremic syndrome. A systemic review, metal-analysis, and meta-regression. JAMA; 290: 1360-1370.

16. George JN, Gilcher RO, Smith JW et al (1998). Thrombotic thrombocytopenic purpura – hemolytic uremic syndrome: diagnosis and management. *J Clin Apheresis*; 13:120-5.

17. Goldwater PN. Bettelheim KA (2000). Escherichia coli 'O' group serology of a haemolytic uraemic syndrome (HUS) epidemic. *Scandinavian Journal of Infectious Diseases*; 32(4):385-94.

18. Grodinsky S, Telmesani A, Robson WL, Fick G, Scott RB (199). Gastrointestinal manifestations of hemolytic uremic syndrome: recognition of pancreatitis. J Pediatr Gastroenterol Nutr; 11(4):518-524.

19. Health Surveillance (2000). The Alberta Oil Sands Community Exposure and Health Effects Assessment Program: Technical Report. Edmonton, Alberta: *Alberta Health and Wellness*. August, 2000. ISBN 0-7785-0225-2

20. Karch H. Bielaszewska M. Bitzan M. Schmidt H (1999). Epidemiology and diagnosis of Shiga toxin-producing Escherichia coli infections. [Review, Tutorial] *Diagnostic Microbiology & Infectious Disease*; 34(3):229-43.

21. Marler Clark. http://www.about-hus.com/. Accessed on Dec. 1, 2003

22. Misselwitz J. Karch H. Bielazewska M. John U. Ringelmann F. Ronnefarth G. Patzer L (2003). Cluster of hemolytic-uremic syndrome caused by Shiga toxin-producing Escherichia coli O26:H11. *Pediatric Infectious Disease Journal*; 22(4):349-54.

23. Nguyen Thu-Ha (2000). Reported Cases of Giardia, Cryptosporidium and Campylobacter in Three Administrative Databases. Banff: SEARCH Program, June 2000.

24. Pradel N. Boukhors K. Bertin Y. Forestier C. Martin C. Livrelli V (2001). Heterogeneity of Shiga toxin-producing Escherichia coli strains isolated from hemolytic-uremic syndrome patients, cattle, and food samples in central France. *Applied & Environmental Microbiology*; 67(6):2460-8.

25. Rogers MF, Rutherford GW, Alexander SR et al. (1986). A population-based study of hemolytic-uremic syndrome in Oregon, 1979-1982. *Am J Epidemiol*;123:137-42.

26. Rowe PC. Orrbine E. Wells GA. McLaine PN (1991). Epidemiology of hemolytic-uremic syndrome in Canadian children from 1986 to 1988. The Canadian Pediatric Kidney Disease Reference Centre.[comment]. J Pediatr; 119(2):218-24.

27. Rowe PC. Orrbine E. Lior H. Wells GA. McLaine PN (1993). Diarrhoea in close contacts as a risk factor for childhood haemolytic uraemic syndrome. The CPKDRC co-investigators. Epidemiology & Infection. 110(1):9-16.

28. Rowe PC. Orrbine E. Lior H. Wells GA. Yetisir E. Clulow M. McLaine PN (1998). Risk of hemolytic uremic syndrome after sporadic Escherichia coli O157:H7 infection: results of a Canadian collaborative study. Investigators of the Canadian Pediatric Kidney Disease Research Center [see comments]. J Pediatr 132(5):777-82.

29. Siegler RL. Postdiarrheal Shiga toxin-mediated hemolytic uremic syndrome (2003). [Comment. Editorial]. *JAMA; 290(10):1379-81*.

30. Siegler RL (1994). Spectrum of extrarenal involvement in postdiarrheal hemolytic-uremic syndrome. *J Pediatr* 125; 511-518.

31. Siegler RL. Griffin PM. Barrett TJ. Strockbine NA (1993). Recurrent hemolytic uremic syndrome secondary to Escherichia coli 0157:H7 infection. *Pediatrics; 91(3):666-8.*

32. Siegler RL. Milligan MK. Burningham TH. Christofferson RD. Chang SY. Jorde LB (1991). Long-term outcome and prognostic indicators in the hemolytic-uremic syndrome. *J Pediatrics; 118(2):195-200.*

33. Siegler RL. Pavia AT. Cook JB (1997). Hemolytic-uremic syndrome in adolescents. *Archives of Pediatrics & Adolescent Medicine; 151(2):165-9.*

34. Siegler RL. Pavia AT. Hansen FL. Christofferson RD. Cook JB (1996). Atypical hemolytic-uremic syndrome: a comparison with postdiarrheal disease. *J Pediatrics;128(4):505-11.*

35. Siegler RL. Pavia AT. Christofferson RD. Milligan MK (1994). A 20-year population-based study of postdiarrheal hemolytic uremic syndrome in Utah. *Pediatrics; 94(1):35-40.*

36. Sorensen H, Sabroe S, Olsen JA (1996). framework for evaluation of secondary data sources for epidemiological research. *Int J Epidemiol*, 25: 435-442.

37. Steinberg EP, Whittle J, Anderson GF (1990). Impact of claims data research on clinical practice. *Int J Technol Assess Health Care*; 6: 282-87.

38. Swerdlow DL. Woodruff BA. Brady RC. Griffin PM. Tippen S. Donnell HD Jr. Geldreich E. Payne BJ. Meyer A Jr. Wells JG (1992). A waterborne outbreak in Missouri of Escherichia coli O157:H7 associated with bloody diarrhea and death. Annals of Internal Medicine; 117(10):812-9.

39. Todd WT. Dundas S (2001). The management of VTEC O157 infection. [Review, Tutorial] *International Journal of Food Microbiology*; 66(1-2):103-10.

40. Wang FL, Gabos S, Shaw S, Xiong X, Schopflocher D (1999). Sources of Diagnosis in Administrative Data: Considerations in Epidemiologic Research and Surveillance. *Area-Wide Epidemiologic and Biostatistics (A-Web) meeting*, Edmonton, April 1999.

41. Wang FL, Semchuk KM, Love EJ (1994). Reliability of environmental and occupational exposure data provided by surrogate respondents in a case-control study of Parkinson's disease. J Clin Epidemiol; 4: 797-807.

42. Wang FL (2005). Using Administrative Data for Research and Surveillance in Alberta: Congenital Anomalies as an Example. Presented at Seminar Rounds, Community Health Sciences, University of Calgary. Calgary, April 29, 2005.

43. Yao P, Wiggs BR, Gregor C, Sigurnjak R, Dodek P (1999). Discordance between physicians and coders in assignment of diagnoses. *Int J Quality in Health Care*; 11:147-153.

APPENDICES

Appendix A. Calculation of Validity Measures

Table Layout for Sensitivity, Specificity, Predictive Value, and Agreement Analysis

		Cases or Events from the Standard		Total
		Yes	*No*	
Cases or Events Being	*Yes*	a (90)	b (15)	**a + b**
Classified from the Study	*No*	c (10)	d (25)	**c + d**
Total		**a + c**	**b + d**	**a+b+c+d = N**

Sensitivity = a / (a + c) * 100 = 90 / (90 + 10) * 100 = 90.0%
Positive Predictive Value(PPV)= a / (a + b) * 100 = 90 / (90 + 15) * 100 = 85.7%
Overall Agreement = (a + d) / N * 100 = (90 + 25) / 140 * 100 = 82.1%
Kappa = P_0 - P_e / 1 - P_e = [2(ad - bc)] /{[(a + b)(b + d)] + [(a + c)(c + d)]}
         = 2(90*25 - 15*10) / (105*40 + 100*35) = 4200 / 7700 = 0.54
         where P_0 is the observed agreement and P_e is the expected agreement.

Appendix B HUS Cases with Lab Confirmation by Year of Reporting

HUS Cases by Lab Testing in the CDRS, Alberta, 1994-2003

Year of reporting	Cases reported in CDRS	Lab Confirmation N	%
1994	9	6	66.7
1995	5	4	80.0
1996	6	4	66.7
1997	6	3	50.0
1998	2	1	50.0
1999	1	1	100.0
2000	6	3	50.0
2001	5	3	60.0
2002	9	7	77.8
2003	9	4	44.4
1994-2003	58	36	62.1

Appendix C. Sex and Age Distribution of HUS Cases by Source of Diagnosis

Sex and Age Distribution of HUS by Source of Diagnosis, Alberta, 1994-2003

Sex Group	Age Group (year)	1st Diagnosis N	1st Diagnosis %	2nd Diagnosis N	2nd Diagnosis %	3rd Diagnosis N	3rd Diagnosis %	All Three Sources N	All Three Sources %
	0-14	76	82.6	10	33.3	5	23.8	91	63.6
Female	15-64	15	16.3	13	43.3	14	66.7	42	29.4
	65+	1	1.1	7	23.3	2	9.5	10	7.0
	All Ages	92	100.0	30	100.0	21	100.0	143	100.0
	0-14	53	84.1	10	52.6	9	69.2	72	75.8
Male	15-64	7	11.1	7	36.8	2	15.4	16	16.8
	65+	3	4.8	2	10.5	2	15.4	7	7.4
	All Ages	63	100.0	19	100.0	13	100.0	95	100.0
	0-14	129	83.2	20	40.8	14	41.2	163	68.5
Both Sex	15-64	22	14.2	20	40.8	16	47.1	58	24.4
	65+	4	2.6	9	18.4	4	11.8	17	7.1
	All Ages	155	100.0	49	100.0	34	100.0	238	100.0

Appendix D. Age Distribution of HUS Cases by Health Region, CIHI data, 1994-2003

Age distribution of newly reported HUS cases by health region, CIHI data, 1994-2003

Residence RHA	Aged 0-14		Aged 15-64		Aged 65 or Over		Total Cases
	N	%	N	%	N	%	
01 - Chinook	21	91.3	1	4.3	1	4.3	23
02 - Palliser	7	63.6	4	36.4	.	.	11
03 - Calgary	48	72.7	15	22.7	3	4.5	66
04 - David Thompson	12	54.5	6	27.3	4	18.2	22
05 - East Central	2	40.0	1	20.0	2	40.0	5
06 - Capital	42	61.8	20	29.4	6	8.8	68
07 - Aspen	12	75.0	4	25.0	.	.	16
08 - Peace County	4	50.0	4	50.0	.	.	8
09 - Northern Lights	4	80.0	1	20.0	.	.	5
Unknown RHA	11	78.6	2	14.3	1	7.1	14
00 - ALBERTA	163	68.5	58	24.4	17	7.1	238

Appendix E. Distribution of HUS Cases by Reporting Hospital and Study Group, CIHI Data, 1994-1999

**Distribution of Reported HUS Cases by Reporting Hospital in Matched and Non-Matched Groups
Alberta, CIHI Data, 1994-1999**

Hospital/Facility Name	Cases Reported in the Matched Group		Cases Reported in the Non-Matched Group		Total Reported Cases
	N	%	N	%	
RAYMOND GENERAL	0	0.0	1	100.0	1
MEDICINE HAT REGIONAL	0	0.0	1	100.0	1
LETHBRIDGE REGIONAL	6	75.0	2	25.0	8
CLARESHOLM GENERAL	0	0.0	1	100.0	1
CALGARY, ALBERTA CHILDRENS	15	55.6	12	44.4	27
CALGARY, FOOTHILLS PROVINCIAL	0	0.0	13	100.0	13
CALGARY, ROCKYVIEW GENERAL	2	40.0	3	60.0	5
CALGARY, GENERAL HOSPITAL GROUP	0	0.0	2	100.0	2
DRUMHELLER REGIONAL HEALTH CENTRE	1	50.0	1	50.0	2
RED DEER GENERAL	1	25.0	3	75.0	4
CORONATION MUNICIPAL	0	0.0	1	100.0	1
Subtotal	**25**	**38.5**	**40**	**61.5**	**65**
EDMONTON, MISERICORDIA	0	0.0	4	100.0	4
EDMONTON, GREY NUNS	0	0.0	4	100.0	4
EDMONTON, ROYAL ALEXANDRA	1	20.0	4	80.0	5
EDMONTON, UNIVERSITY OF ALBERTA	0	0.0	29	100.0	29
MAYERTHORPE GENERAL	0	0.0	1	100.0	1
BARRHEAD GENERAL	0	0.0	1	100.0	1
ATHABASCA GENERAL	0	0.0	1	100.0	1
VEGREVILLE, ST. JOSEPH'S GENERAL	0	0.0	1	100.0	1
FOX CREEK HOSPITAL	0	0.0	1	100.0	1
FAIRVIEW GENERAL	0	0.0	1	100.0	1
GRANDE PRAIRIE, QUEEN ELIZABETH II	0	0.0	2	100.0	2
WABASCA/DESMARAIS GENERAL	0	0.0	1	100.0	1
FORT MCMURRAY REGIONAL	0	0.0	1	100.0	1
Subtotal	**1**	**1.9**	**51**	**98.1**	**52**
ALL REPORTING HOSPITALS	26	22.2	90	76.9	117

Appendix F. Methodological Considerations in Case Definition

Case definition is an important but complex issue in using administrative data. How to define a case has a direct impact on validity and completeness of data and on case count and rate estimation. A case may be identified by using diagnostic codes and/or service procedure codes available in the system.

The first thing to be considered is how many diagnoses should be used in "case ascertainment". Traditionally, health reports and studies used only primary diagnosis in case ascertainment. However, as identified by our previous work (Wang et al., 1999) and the current study, a proportion of cases may be missed if only the primary diagnosis is used. The extent of missed proportion will depend on the complexity in clinical diagnosis, nature (acute vs. chronic, disease vs. syndrome) and severity of a condition, patterns of seeking care among population groups, policy of health care services, patterns of practice and coding, and process of reporting, etc. For instance, diagnosis of a cleft lip is easier than diagnosis of a heart defect. This could lead to, relative to a heart defect, a better validity and completeness of cases with cleft clip reported in the CIHI Hospital Inpatient System if other factors hold the same for both diseases. However, other factors are likely different in reality. Heart defects are clinically more serious than cleft lip, thus more likely being admitted into hospital for treatment, and being coded as primary diagnosis and reported.

HUS is a rare but severe syndrome that may result from other diseases. Therefore, a patient is more likely hospitalized but may not be necessarily coded as HUS in the primary diagnosis. In the current study, about 35% of HUS cases were captured in 2nd and 3rd diagnoses. To capture all potential cases in a study, ideally all diagnoses available need to be considered. In practice however we need to balance the "trade off" between the sensitivity and the specificity, and between the sensitivity and the complexity of data management and analysis. While using more diagnoses may increase the sensitivity of a case identification procedure it may decrease the specificity at the same time, thus sacrificing the accuracy of diagnosis and increasing complexity of data management and analysis. For monitoring purposes and regular surveillance activities, the first three diagnoses perhaps are sufficient for most diseases.

The 2nd issue to be considered in case ascertainment and definition is how many data systems should be used. Although there may be overlaps for a same patient across data systems, such as Fee-for-service (FFS) Claims, Ambulatory Care System (outpatients, including emergency visits), and Hospital Inpatient System, a proportion of patients of many diseases and conditions may be captured only by one of the data systems. For instance, in children 0-17 years only about 53% of congenital anomalies are captured by the Hospital Inpatient System while the remaining 47% are captured by FFS Claims and/or Ambulatory System in 1998-2003 (Wang, 2005). Given the fact that the administrative data systems only capture the patients who seek for care – thus likely to underestimate the incidence and prevalence, using all three data systems available is recommended to maximize the case ascertainment from administrative data systems.

There are other factors that require consideration in case definition, such as health care service procedure, prescription code for drug use, sex/age distribution of disease, general frequency of health care visit in a given time period, the interval between the visits, hospital admission, coding practice, etc. Ideally, a case definition should be constructed while taking all of these factors into consideration. In the present study, the procedure code was used to cross check the diagnostic code. Findings of relevant treatment to HUS increase, to some degree, confidence in the validity of our case definition. Similarly, uneven female/male ratio across the age groups and uneven age distribution across the three diagnosis groups may warrant attention in the development of HUS case definition using administrative data. Sometimes, a service procedure code/prescription code is very specific to a given disease/condition, such as breech delivery (ICD-9 = 72.5) for pregnancy and antidiabetics agents (code = 682008) for diabetes, thus can be used for case identification and definition. The frequency of visits and interval between visits were used in case definition of asthma in Oil Sands study (Health Surveillance, 2000) and appeared to be very useful. The policy of hospital admission and the practice of patient coding are factors that can also affect the case ascertainment and reporting.