uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES
ET POSTOCTORALES

u Ottawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND
POSDOCTORAL STUDIES

Wenxia Zhao
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Epidemiology)
GRADE / DEGREE

Department of Epidemiology and Community Medicine
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Comorbidity in Prediction of In-hospital Mortality Among Diabetic Patients :
A Study-derived Index

TITRE DE LA THÈSE / TITLE OF THESIS

Yue Chen
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Bernard Choi

George Wells

Gary W. Slater
LE DOYEN DE LA FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES /
DEAN OF THE FACULTY OF GRADUATE AND POSTDOCORAL STUDIES

COMORBIDITY IN PREDICTION OF IN-HOSPITAL MORTALITY AMONG DIABETIC PATIENTS

- A STUDY-DERIVED INDEX

BY

WENXIA (HELEN) ZHAO

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements for the degree of

Master of Science

In

Epidemiology

Department of Epidemiology and Community Medicine
Faculty of Medicine
University of Ottawa

March, 2005

Your file Votre référence
ISBN: 0-494-11475-4
Our file Notre référence
ISBN: 0-494-11475-4

ABSTRACT

The present study developed and validated a comorbidity index specifically for prediction of in-hospital mortality among diabetic inpatients in Canada. The analysis was based on data from the Hospital Person-Oriented Information Database (HPOI) for the study period from 1995/96 through 2000/01. The study included all the hospitalizations with a primary or secondary diagnosis of diabetes (ICD-9 code: 250.x) in acute care hospitals for patients aged 45 years or older with a length of stay of 90 days or less in ten provinces. All episodes of hospitalization for each patient were linked using a unique patient identifier, and one was randomly selected for the analyses. The study population of 578,222 diabetic inpatients was randomly divided into two parts, which were used either to develop or to validate the index. Multiple logistic regression models were used to develop and validate the index. A total of 22 diabetic comorbidities including 14 coexistent general medical conditions and 8 diabetic complications were included in the study-derived index, which had a better predictive performance as compared with D'Hoore-Charlson index and the simple count of comorbidities. The study-derived index can be used to control for potential confounding caused by comorbidity in the exploratory data analysis of diabetes research, to assist in creating more effective diabetes management system and to evaluate the prognosis of diabetic inpatients for health care provider.

ACKNOWLEDGEMENTS

My foremost thank goes to my thesis committee, Dr. Yue Chen, Dr. Ronald J. Sigal and Dr. Helen Johansen. Without them, this thesis would not have been possible. I thank them for their insights, guidance and direction throughout the entire project, and for their patience and encouragement that carried me on through difficult times. Their valuable feedback contributed greatly to this thesis.

I thank Health Statistics Division at Statistics Canada for the approval of my access to the database and the pleasing work environment. I thankfully acknowledge the financial support provided by Canadian Institute for Health Research in the first year of this thesis project.

I thank the Department of Epidemiology and Community Medicine at University of Ottawa for the opportunity to pursue my Master's degree. A special thank to Fay Draper for her understanding, time and great help, which make my study at University of Ottawa a pleasing experience.

I would like to give my thanks to my colleagues and friends. To Kathy Nguyen for her great help on SAS programming at the beginning of this project. To Mei Lin and Vanita Economou for sharing their ideas with me and making this learning experience enjoyable. To Christie Sambell for her help in proofreading the thesis. To Nawaf Madi for his understanding and support.

Finally, I would like to thank my parents for their unconditional love, encouragement and support. I thank my sister and brother for encouraging me to pursue the Master's degree, which I once thought was not possible. Especially, I thank my husband for always being there for me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Diabetes mellitus affects more than 2 million Canadians, and approximately 30,000 deaths each year are attributed to diabetes and related complications (1). In view of its associated morbidity, mortality and costs, diabetes mellitus constitutes a major health care burden among chronic illnesses (1). Patients with type 2 diabetes mellitus often develop chronic micro- and macrovascular complications, of which end-stage renal disease, cardiovascular disease and cerebrovascular disease are the major contributors to the excess mortality associated with diabetes (2). Approximately 21% of people with diabetes compared with 4% without diabetes have heart disease or are suffering the effects of stroke (3). People with diabetes also appear to be at increased risk for general medical conditions other than acute glycaemic and chronic complications of diabetes (4).

Diabetes mellitus is a substantial cause of mortality at all ages (5), and the mortality in adults with type 2 diabetes is at least twice as high as in the general population (6). In 1997, diabetes contributed to more than 190,000 deaths in the United States, and was listed as the underlying cause of death for approximately 62,000 people, making it the seventh-leading cause of death (7). Life expectancy of middle-aged persons with type 2 diabetes is reduced by 5 to 10 years on an average, and the reduced life expectancy is greater for individual with complications (6). In 1998, the value of lost production due to premature mortality accounted for 45 % of the total economic burden of diabetes, which represents indirect cost of $732.8 million in Canada (8).

Due to the lack of a solid basis to differentiate between diabetic complications and coexistent general medical conditions, a comorbidity of diabetes can be broadly defined as any distinct additional clinical entity that has existed or that may occur during

the clinical course of a patient with diabetes (9-11). People with diabetes are more likely to be hospitalized for a comorbidity than people without diabetes. Compared to their nondiabetic peers, middle-aged persons with diabetes were at 60% greater risk of hospitalization for general medical conditions overall (Relative Risk (RR): 1.6; 95% CI: 1.2-2.0); elderly persons with diabetes were at increased risk of hospitalization for selected general medical conditions, such as liver disease (RR: 3.0; 95% CI: 2.9-3.0) and septicemia (RR: 2.8; 95% CI: 2.8-2.9) (4). In 1997, of all hospitalizations attributable to diabetes in the U.S., the six chronic complications of diabetes including neurological disease, peripheral vascular disease, cardiovascular disease, renal disease, ophthalmic disease and other chronic complications accounted for 35.0%; coexistent general medical conditions such as liver disease, respiratory failure, malignant neoplasms and affective disorders accounted for 51.6%; and uncomplicated diabetes and its acute metabolic complications only accounted for the remaining thirteen percent (13.3%) (12).

Comorbidity is an important predictor of mortality for people with diabetes. In a study of 1,779,167 adult and nonmaternal inpatients from 438 acute care hospitals in California in 1992, it was found that the greater the number of comorbidities affecting a patient, the greater the likelihood of in-hospital death. Death in hospital was seven times as high for patients with three or more comorbidities compared with patients who had no comorbidities (13). A study conducted in Spain indicated that the presence of comorbidity was associated with an increased risk of dying in the inpatient period among diabetic patients (odds ratio=3.4; $p<0.01$) (14). In a study of seven hospitals, Greenfield et al found that the hospitals with higher average degree of comorbidity had an increased risk of mortality (15).

A comorbidity index can estimate the impact of comorbidity on health care outcomes, and is a potentially useful tool for health-service researchers, health policy makers and clinicians. Comorbidity indices are increasingly used in health service research. Based on medical record review, Charlson and colleagues developed a weighted-index measure of comorbidity that was shown to predict one-year all-cause mortality (16). The Charlson comorbidity index and its adapted versions for use with administrative data, such as the Dartmouth-Manitoba, Deyo and D'Hoore-comorbidity indices, have been widely used in health research. There are some other comorbidity indices, such as chronic disease scores derived from pharmacy databases (17; 18) and an index derived from physician claims data (19). However, all the existing comorbidity indices were derived from some specific populations such as patients who underwent surgery, elderly people or patients with cancers, many of whom are different clinically from many people with diabetes. The Charlson Index was derived from a narrowly defined clinical population of fewer than 600 patients and only included comorbidities that happened to occur in this specific small population (13). Some other comorbidities not included in the Charlson index might also be predictive of mortality (20) or relevant to patients with a particular condition or procedure (21). There is no previously published comorbidity index for diabetic patients, and little is known about the relative performance of the existing comorbidity indices in predicting in-hospital mortality among diabetic patients. Therefore, the existing comorbidity indices may include some diseases that may have no significant impact on clinical outcomes for people with diabetes and excluded other diseases, which may be relevant to outcomes in diabetic patients.

The review by Schneeweiss et al (22) found that study-specific weights improved the validity of the Charlson index measure of comorbidity. Many authors have recommended that researchers working with large scale administrative databases should use multivariate analysis to derive study-specific comorbidity weights for different populations and different outcomes (20; 21; 23; 24). This creates a comorbidity index that is "tailored" to the patient population, condition (surgical or medical), and outcomes under study (20), which may provide the best method of risk adjustment (9; 20).

2. OBJECTIVE

- To develop and validate a new comorbidity index specifically for prediction of in-hospital mortality among diabetic inpatients in Canada.
- To compare the new index to the D'Hoore-Charlson index (Appendix 1) and a simple count of the number of comorbidities in the prediction of in-hospital mortality as well as the odds of a longer stay at hospitals in diabetic inpatients.
- To describe the importance of comorbid conditions in the prediction of in-hospital mortality among age-specific diabetic inpatient populations.

3. LITERATURE REVIEW

3.1. Definition of Comorbidity

The definition of comorbidity is not consistent across studies. In some studies, comorbidities are narrowly defined as coexisting medical conditions: 1) that are distinct from the principal diagnosis or the primary illness for which the patient seeks health care services (11; 19; 25); and 2) are not causally related to the principal disease process (9;

13; 26; 27). The study populations in these cited studies are usually patients with a uniform principal diagnosis, such as patients who undergo lumbar spine surgery. In these studies, comorbidities refer to the total burden of illnesses that are unrelated to the patient's principal diagnosis or treatments (25) and have important implications regarding patient outcomes (27). In contrast to complications, a narrowly defined comorbid condition is not linked causally to the natural history or treatment of the principal diagnosis (26). For example, for a patient admitted to hospital undergoing prostatectomy, local infection would be considered as a complication of the surgery and diabetes would be considered as a comorbid condition (28). Differentiating pre-existing comorbidities from complications of treatment or health care is a key point for the studies evaluating the effectiveness of treatments.

A broadly defined comorbidity refers to any distinct additional clinical entity that has existed or that may occur during the clinical course of a patient who has the index disease under study (11; 29). Under this definition, comorbidities include coexistent conditions that are causally related (complications of an index disease) and not related (general medical conditions) to an index disease (9-11; 30), and the study populations with index disease usually have a wide range of principal reasons for hospitalization. Comorbidity indices developed or used in these studies include principal diagnoses, complications of the index disease, and general medical conditions. Therefore, they represent a total morbidity measure in this specific study population (9). For a chronic index disease like diabetes, complications of the index disease and its coexistent general medical conditions together increase the burden on health care and worsen the health care outcomes. Therefore, the broadly defined comorbidity may provide a more complete

picture of disease factors that influence health care burden and outcomes among people with an index chronic disease. In addition, due to the limitations of administrative data, most studies cannot distinguish pre-existing comorbid conditions from complications of index diseases.

As a whole, comorbidity has important implications for clinical outcomes, including mortality, surgical results, functional status, and length of hospital stay, as well as economic outcomes, including resource utilization, discharge destination, and intensity of treatment (9). Researchers should choose either broadly or narrowly defined definition of comorbidity according to their research questions.

3.2. Existing Comorbidity Indices

In 1974, Kaplan and Feinstein introduced a comorbidity scoring system that classifies diabetic patients from grade 0 (no comorbidity) to grade 3 (severe comorbidity). The scheme is based on consensual clinical judgment of the relative importance of a variety of comorbid conditions (20). The Charlson comorbidity index (16) was originally developed to predict one-year all-cause mortality in 1987 using data abstracted from the medical records of a cohort of 559 medical patients. It assigns weights of 1, 2, 3 and 6 for each of the existing comorbid diseases to derive a total score (16), which represent a measure of the burden of comorbid disease (31). It provides a prospectively applicable method for classifying comorbid conditions that might alter the risk of mortality for use in longitudinal studies (16). After the development of the Charlson comorbidity index, some independent investigators have developed modified versions of Charlson index to adapt it for use with administrative databases, and the commonly used indices include Dartmouth-Manitoba Charlson index by Roos *et al.* (21; 24; 32; 33), Deyo-Charlson

index (34), D'Hoore-Charlson index by D'Hoore *et al.* (31; 35) and Ghali-Charlson index (20). Deyo *et al.* (34) , the Dartmouth-Manitoba group (21; 24; 32; 33) and Ghali et al. independently assigned ICD-9-CM codes to the diagnoses in the Charlson Index. While Dartmouth-Manitoba Charlson index and Deyo-Charlson index differ slightly in their selection of ICD-9-CM codes to represent the various comorbidities (36; 37) and use the original Charlson index weights, Ghali et al. reduce the number of diagnoses by only selecting the significant predictors to their outcome and changed the weights to improve the predictive performance (20). Motivated by the fact that some institutions, frequently outside the US, use only ICD-9 codes without the Clinical Modification (CM), and coding of the tailing digits in the ICD-9-CM is less reliable, D'Hoore *et al.* modified the Charlson Index by using only the first three digits of ICD-9 (22).

Among the different systems developed, the Charlson Index (16) offers a series of strong points: 1) it is relatively easy to construct; 2) its versions for administrative databases based on the International Classification for Diseases are available; 3) and it is widely used (38). Although it was developed for the purpose of prospectively predicting one-year mortality, it has been applied to discharge abstract and claims data to predict short-term outcomes such as in-hospital mortality, blood transfusions, hospitalization charges, and length of stay (19-21; 24; 31-35).

Various other risk adjustment indices have been developed to measure the burden of comorbidities. Two versions of the Chronic Disease Score use pharmacy dispensing data to assign patients to chronic disease groups (22). A person's Chronic Disease Score is the sum of the weights corresponding to the different medication classes, regardless of how many different medications he or she is taking within a given class (39). As a

measure of comorbidity, the Chronic Disease Score is appealing because it is simple, easy to use and inexpensive, and relies on information that is readily available (39). However, using drug prescriptions as proxies for diagnoses has uncertain validity because: 1) medication often have mixed indications (22; 40; 41); 2) physicians tend to avoid prescribing additional drugs to patients who are already taking several of them and thereby unfortunately may underprescribe preventive medication in sicker patients (22; 40; 41); 3) drug use varies not only by disease status but also by ability to pay, prescribing customs and patient attitudes, which vary among health systems and regions (22). These limitations might be the reasons that they are not commonly used to develop comorbidity index.

Some researchers found that an index that simply counts the number of diagnoses provides good explanatory ability in models assessing outcomes, and performs almost as well in prediction of mortality as the more complex Charlson scores (42; 43). It is suggested that Charlson scores may be more susceptible to coding errors compared to simple counts of the number diagnoses (43). Elixhauser et al. adjusted for separate comorbidity diagnoses instead of combining comorbid conditions into a single summary score (13). Their system might control confounding slightly better, however, it is not always feasible because a regression model that uses so many variables to control for comorbidity limits the ability to model interactions and decreases precision in epidemiological analyses. Furthermore, it can only be performed in rather large population (22). An index incorporating a very large number of variables, each with its own coefficient, would also be cumbersome to use in clinical practice.

Table 1 lists some comorbidity measuring instruments.

Table 1: Comorbidity measuring instruments

Instrument	Authors	Publication Year	Settings of original study	Information used	Population under study	Prognostic endpoint	Statistical model used
Original Charlson Index (16)	Charlson et al.	1987	1) Development*: New York Hospital-Cornell Medical Center during 1-month period in 1984; 2) Test*: Yale New Haven Hospital between Jan.1, 1962 – Dec.31, 1969	Medical records	1) Development: 559 medical patients; 2) Test: 685 women with breast cancer	One-year mortality	Cox proportional hazard regression model
Charlson Index in versions for administrative database							
Deyo-Charlson Index (34)	Deyo et al.	1993	Medicare beneficiaries in the USA in 1989	Medicare claim database: ICD-9-CM coded diagnoses	27,111 patients who underwent lumber spine surgery with a mean age of 72 years	Post-op death (in-hospital or 6 weeks after discharge), post-op complications, length of stay, hospital charges	Logistic regression, Multiple linear regression

Instrument	Authors	Publication Year	Settings of original study	Information used	Population under study	Prognostic endpoint	Statistical model used
DM-Charlson Index (21; 24; 32; 33)	Roos and Romano et al.	1989/1993/1997	1) A Manitoba hospital in Canada, April 1980-May 1992; 2) Licensed federal hospitals in California, USA in 1988-1991	Hospital discharge database: up to 16 and 25 ICD-9-CM coded diagnoses in studies 1) and 2) respectively	1) 4121 patients having bypass surgery aged ≥30 years 2) 55 407 patients having lumbar disc excision surgery aged ≥18 years	One-year mortality, significant in-hospital complications, and 90-day readmission	Logistic regression
D'Hoore-Charlson Index (31; 35)	D'Hoore et al.	1993/1996	792,839 records in 78 hospitals with > 100 beds in Quebec, Canada in 1989/90	MED-ECHO database: ICD-9 coded up to 16 diagnoses	33,940 inpatients with ischaemic heart disease with a mean age of 63 years	In-hospital mortality	Logistic regression
Ghali-Charlson Index (20)	Ghali et al	1996	1) Development: 257,333 discharges in 1990; 2) Test: 267,407 discharges in 1992 from all Massachusetts (USA) hospitals	Massachusetts Health Data Consortium discharge abstracts: up to 15 ICD-9-CM coded diagnoses	1) Development: 6326 patients in 1990; 2) Test: 6791 patients in 1992 who underwent bypass surgery with a mean age of 65 years	In-hospital mortality	Logistic regression

Instrument	Authors	Publication Year	Settings of original study	Information used	Population under study	Prognostic endpoint	Statistical model used
Chronic Disease Scores (CDS) (18)	Von Korff et al.	1992	Enrollees of the Group Health Cooperative of Puget Sound in Seattle and surrounding communities in 1985	Pharmacy database: information on all prescription medication	122,911 adult enrollees in Central and East region	One-year mortality, one-year hospitalization rate	Logistic regression
Extended Chronic Disease Scores (CDS-2) (17)	Clark et al.	1995	Enrollees of the Group Health Cooperative of Puget Sound in Seattle and surrounding communities in 1992	Pharmacy database: information on all prescription medication	1) Development: randomly 50% of adult enrollees: 125,000. 2) test: second half of adult enrollees	One-year primary care visits, one-year total costs, one-year outpatient costs	Logistic regression
Physician claims comorbidity index (19)	Klabunde et al.	2000	National cohort during the period January 1, 1992 through December 31, 1993	The National Cancer Institute's Surveillance, Epidemiology and End Results (SEER) program and Medicare claims (ICD-9-CM)	Elderly prostate (n = 28,868) and breast cancer (n=14,943) patients	Two-year non-cancer mortality	Cox proportional hazard model

Instrument	Authors	Publication Year	Settings of original study	Information used	Population under study	Prognostic endpoint	Statistical model used
Sum of number of comorbidities (42; 44-46)	Siegler et al. (Used it)	1994	Rehabilitation centers in U.S.	retrospective cohort data	1,075 medical rehabilitation patients	acute medical complications on rehabilitation requiring transfer	Logistic regression
Kaplan and Feinstein index (47)	Kaplan MH, Feinstein AR	1974	The West Haven VA Hospital during the years 1959-1962	Data abstracted from medical records	615 Peoples with diabetes	5-year mortality	consensual clinical judgment, and no statistical modes used
APACHE II (Acute Physiology and Chronic Health Evaluation (48; 49)	Knaus et al.	1985	Intensive care units at 13 US hospitals during 1979-1981	Data abstracted from medical records	5,815 intensive care admission	Death in hospital	Logistic regression
APACHE III (Acute Physiology and Chronic Health Evaluation (48; 49)	Knaus et al.	1991	Intensive care units at 40 US hospitals during 1979-1981	Data abstracted from medical records	17,440 unselected adult medical/surgical intensive care unit admission	Death in hospital	Logistic regression

* "development" represent the sample used to develop an index; "test" represent the sample used to test an index.

3.3. Data Sources

Although many risk adjustment instruments have been developed to measure the burden of comorbidity, there are mainly two approaches to adjusting health care outcomes for clinical risk: systems using large scale administrative data sets that contain only limited clinical information and systems that abstract detailed clinical information from the medical record (50). Administrative data as a source of comorbidity study have some advantages and disadvantages as compared to medical records.

3.3.1. Medical record data

Some investigators found that medical record data are superior to administrative data for predicting mortality (37; 51) because it allows abstraction of detailed clinical information. However, the following disadvantages of medical record data limit its use in developing comorbidity scores: 1) extracting clinical detail for more precise measurement is prohibitively expensive on the scale needed to develop a general comorbidity tool (13); 2) the results obtained from medical records may not be broadly applicable to the administrative data that are in common use in health research (13); 3) sample sizes are usually small; and 4) accessibility of the required range of medical charts may be limited.

3.3.2. Administrative data

Advantages include that: 1) administrative data provide large, diverse, population-based, and representative samples that are appropriate for studying real-world difference in patient outcomes (9; 21); 2) the cost of administrative data is much lower, and the

required data collection time is reduced (37); and 3) administrative data are ready to use and easy to access (21; 52).

There are also several disadvantages. First, administrative data tend to underestimate the presence of comorbid conditions (22; 24; 26; 37; 45; 53; 54). The study by Romano et al found that administrative data revealed few patients with composite scores greater than one, and detailed chart reviews shifted the distribution of the comorbidity index significantly toward higher scores (21). This study also supports Charlson's contention that differences in how comorbidities are defined or coded may contribute to differences in the performance of comorbidity indices based on clinical and administrative data (21; 55). Misclassifying patients who truly have a comorbidity may bias the estimate of risk either upward or downward, depending on whether the misclassification is differential or nondifferential. A prominent example is that controversial conclusions have been obtained from discharge information and medical record review. Roos *et al.* found that transurethral prostatectomy appeared to be associated with increased 5-year mortality after comorbidity adjustment based on discharge information (RR = 1.45, 95% CI: 1.15–1.83) (56; 57). Concato *et al.* repeated the study and found similarly elevated risks when using administrative data. However, the increase in mortality vanished (RR = 1.03, 95% CI: 0.51–2.07) when the same method was used based on medical record review (58).

Second, many comorbidity studies found a bias in discharge abstract coding – a recording bias that reduces the likelihood of chronic unthreatening comorbid conditions being reported for patients who are severely ill or die (19; 25; 50; 59-64). Consequently, these comorbid conditions are spuriously associated with a lower risk of in-hospital

mortality, which is clinical implausible (13; 19; 25; 62-64). The two possible reasons for a recording bias may be: 1) physicians are less likely to record chronic conditions for patients who are severely ill or die (13; 25; 27; 50); and 2) a seriously ill patient may have had so many medical problems that non-threatening chronic diseases, although detected, are discarded because the data fields have been exhausted by more important diagnoses (22; 65; 66). Romano *et al.* showed that sensitivity to capture specific diagnoses in administrative databases with five diagnosis fields was reduced by an average 13% point compared with a record with 25 fields, and comorbidity is much better reported when the claims is extended to 9 or 24 diagnoses (66).

Third, while discharge abstract data typically contain discharge diagnoses reflecting all conditions addressed during hospital stay, they do not indicate whether a given diagnosis was present on admission or arose later in the hospitalization (13; 22; 25; 67-69). Discharge abstract data also does not distinguish between independent conditions and conditions that are directly related to the principal diagnosis (13).

Fourth, the distinction between the principal diagnosis and secondary diagnoses may be arbitrary and based on nonclinical decisions in discharge abstract data (13; 25). Although there are clear coding guidelines for what constitutes a principal diagnosis, these guidelines may be subordinated to other incentives such as the desire to maximize reimbursement (13; 25). The last two shortcomings of discharge abstract data may have important influences on the studies using the narrow definition of comorbidity, but not on the studies in which the broad definition of comorbidity is used.

3.3.3. Is administrative data a proper source for developing an index?

The success of development and usefulness of a comorbidity index varies and depends not only on the statistical methodology employed, but more importantly, on the completeness and accuracy of available data (23). There is a concern about the value of the index derived from discharge abstract data because diagnoses are occasionally erroneous, often coded incorrectly, and frequently omitted from administrative data (22; 50). One research recommended that great caution be taken in using discharge diagnoses of comorbid conditions to adjust hospital death rates for clinical differences in the patients populations (50).

However, comorbidity indices involving special data abstraction from medical records also impose a significant burden of data collection (50). In addition, they would not be broadly applicable to the administrative data that are in common use today (13). Furthermore, one study indicated that both medical records and administrative data are likely to misclassify comorbidity to some extent, and the medical record-based index is not a gold standard, but a source that also measures comorbidity with error (70).

Finally, the results of the studies comparing the predictive abilities of indices derived from medical records and administrative data are controversial. One study found the Charlson comorbidity index derived from medical record data to be superior to the same index derived from administrative data (37). The study by Hannan *et al.* (51) also suggested that a clinical database was superior to an administrative database for predicting in-hospital mortality among CABG patients (51). However, in another study the chart-based index yielded higher scores and identified more conditions than the ICD-9-based index, but the chart-based index did not have superior mortality prediction (9). A

study conducted among patients undergoing prostatectomy indicated that comorbidity as measured by claims data was only slightly less predictive of patient survival than comorbidity as measured by medical record data (45).

Administrative data are being used with increasing frequency by third party payers, managed care companies, and researcher to make management decisions, formulate policy, evaluate medical providers, and assess medical effectiveness and outcomes (23). Chart review is rarely possible in the growing number of low-budget health studies. Therefore, administrative data with all its limitations is still commonly used to develop comorbidity indices (22). Researchers suggest that a comorbidity score derived from administrative data sets is an acceptable first step in controlling for differences across patient populations (26; 70).

3.4. Development of Comorbidity Indices

The original Charlson index was derived from a cohort of 559 patients admitted to an acute hospital during a certain period, and validated in another cohort of 685 women with histologically proven primary breast cancer who were followed up exhaustively over 10 years (16). Many researchers only developed their indices without validating them on independent populations due to the limit of sample size or sources of study populations. In a few studies with populations large enough, modeling was carried out using a randomly selected half of populations (i.e., "development" datasets), and the remaining half of the dataset was employed to assess the indices' predictive ability (19; 70).

3.4.1. Statistical models

For studies based on administrative databases, logistic regression models have been used to assess the contribution of the comorbidities to predicting in-hospital mortality (13; 20; 31; 37; 69; 71) because mortality information in administrative data is usually limited to status at time of discharge (20). Multiple logistic regression models contain a set of independent variables that are commonly used to isolate the influence of demographics, financial incentives, and clinical differences among patients (13; 31; 37; 69). Studies using medical records have more detailed information including survival time. For these studies, Cox proportional hazards models are used to estimate the adjusted hazard of mortality (19; 70).

3.4.2. Weights of comorbidity index

The Charlson index was constructed by awarding a weight to each of the comorbidities (1, 2, 3, or 6) depending on the magnitude of their adjusted relative risk associated with mortality, and calculating an overall score from the sum of weights (16). Such a weighted comorbidity index takes into account both the number and the seriousness of comorbid diseases. The predictive abilities of indices using exact relative risks and relative risks rounded to the nearest integer were compared, and the results did not differ significantly (16). Satariano and Ragland proposed a new weighting system with the aggregate comorbidity score being simply the sum of the number of comorbid conditions present (44). The Kaplan and Feinstein index assign a value of one to three to the comorbidity in each of vital body systems. The score for the most severe condition typically becomes the aggregate comorbidity score. When the highest single system score

is a two, but a score of two is reported in more than one system, a Kaplan and Feinstein summary score of three is assigned (47).

Most studies used the method proposed by Charlson et al (16) to round the adjusted relative risks to the nearest integer. Conditions with relative risk of less than 1.2 were dropped from consideration; conditions with relative risks of 1.2 - 1.4, 1.5 - 2.4 and 2.5 - 3.4 were assigned weights of 1, 2 and 3, respectively. The two conditions having relative risks of 6 or more were assigned a weight of 6.

A few studies used the estimated coefficient values (rather than adjusted relative risk) as weights (19). The use of simple integer weights – rather than coefficients – is associated with a decrease in model c statistics and r values. This is not surprising because the assigned integer weights only approximate the more precise coefficient (20). However, to simplify comorbidity adjustment systems, most investigators still prefer to use simple integer weights.

3.5. Validation of Comorbidity Indices

There is no gold standard for validating or comparing comorbidity indices (22). In early papers, the validity of prediction is often assessed by the strength of an association between the comorbidity index and the outcome, in terms of the odds ratios (OR) or relative risk (RR) per increment in score (21; 22). This method of validation has been regarded as inappropriate because the statistical association between indices and outcomes should not be considered as synonymous of validity and indices should be evaluated on the basis of their power to improve the discriminating capacity in models (38).

Because of the absence of a gold standard, researchers have to assume that 'true comorbidity' is correlated with worse health care outcomes, health care utilization and costs. Therefore, the validity of a comorbidity index is assessed by how well the index predicts those outcomes which indirectly determines how well it can control for confounding (22). For dichotomous endpoints (such as mortality) modeled by fitting logistic regression models, most researchers use *c* statistic (20; 22; 23; 33; 42; 71; 72) and the likelihood ratio chi-square statistic (G^2) (20; 23; 42; 71), and some also use Hosmer-Lemeshow test (13; 20-22; 66) to validate comorbidity indices.

3.5.1. c statistic

The c statistic is a measure of model discrimination (20). Discrimination reflects how well a model distinguishes between individuals with poor outcomes and those with good outcomes, and is commonly quantified using the area under curve (AUC) of a receiver operating characteristic (ROC) (21; 22). As the most valid measure of predictive accuracy, the c statistic represents the probability that a randomly selected pair of subjects, one a survivor, the other a nonsurvivor, is correctly rated (31; 33). The AUC or *c* ranges from 0 to 1 with 1 indicating a perfect prediction and 0.5 indicating a chance prediction. It has been suggested that c statistic of 0.7-0.8 could be considered acceptable and those of 0.8-0.9 excellent (22).

3.5.2. Likelihood ratio chi-square statistic (G^2)

Likelihood ratio chi-square statistic (G^2) assesses whether the addition of a comorbidity index significantly improves the model. Unconditional logistic regression models are used to examine the relative importance and the explanatory power of indices

for predicting outcomes, such as short-term mortality. The contribution of an index is assessed via the likelihood ratio chi-square statistic (G^2) resulting from comparing a model containing age, gender, and the comorbidity index with a nested baseline model that excluded the comorbidity index (23).

Since the likelihood ratio chi-square statistic (G^2) is affected by the number of independent variables in a model, when researchers compared the performance of different indices, they performed additional analyses using each index as a continuous variable. This way, all models had only one more independent variable than the baseline model (42).

3.5.3. Hosmer-Lemeshow test

The Hosmer-Lemeshow goodness-of-fit statistic (H-L) is a measure of calibration (20; 22). It describes the agreement between observed and expected mortality for patients rank-ordered by risk strata: a small value of the Hosmer-Lemeshow chi-square statistic and a correspondingly high p-value suggest a good fit of a model (20; 21; 33). Models with Hosmer-Lemeshow χ^2 values greater than 15.51 (i.e., $p<0.05$) are rejected for poor fit (20). However, in the studies with large sample sizes, the result of Hosmer-Lemeshow test should be cautiously interpreted since even small differences between the observed and expected values become significant (13).

See "Research design and method" for more detail about the above statistics.

4. RESEARCH DESIGN AND METHODS

4.1. Data Source

Our analysis was based on data from the Hospital Person-Oriented Information (HPOI) Database for the study period from 1995/96 through 2000/01, maintained by Statistics Canada. This database contains information on inpatient hospital separations (discharges or deaths) from most acute care and some psychiatric, chronic and rehabilitation hospitals across Canada. Each record contains information for one patient separation (discharge), and a separate record is created for a readmission. Excluding error codes, records for newborn babies, non-resident records, and records without valid personal identification number, approximately 85% of all hospital morbidity records and 95% of all non-newborn records are included in the POI database. The data are based on the April-to-March fiscal year (73).

The HPOI database provides linkages of all hospitalization records for each individual within a province, which allow records to be linked internally (within a year) and externally (between years) by using a unique patient identifier. Each hospitalization record contains the primary diagnosis and 15 other diagnosis codes except for those from the provinces of Prince Edward Island and Saskatchewan. Five small hospitals in the province of Prince Edward Island and all of Saskatchewan only report primary diagnosis and two secondary diagnoses. The first or the tabulation ICD-9 code describes the most responsible diagnosis, which is the main cause of the patient's stay in hospital. Other diagnoses are designated and defined as all conditions that coexist at the time of admission, or that affect the treatment received and/or length of stay (73). Diagnoses are identified using the International Classification of Disease, 9th revision (ICD-9) code. The

information obtained from the HPOI Database included age at admission, sex, province where patients were admitted to hospitals, admission date, separation date, medical diagnoses at discharge containing primary diagnoses and up to 15 other diagnoses, discharge status (alive or dead), and type of hospital (acute or not) or type of service.

4.2. Study Population

All the hospitalizations with a primary or secondary diagnosis of diabetes (ICD-9 code: 250.x) in acute care hospitals from 1995/96 to 2000/01 were chosen from the HPOI datasets. Adult patients 45 years of age or older, who were admitted to acute hospitals with a length of stay of 90 days or less during 1995/96 to 2000/01 in all ten provinces, were included in the study. Patients living in the territories of Yukon, Northwest Territories and Nunavut, who are likely to be transferred to larger hospitals elsewhere, were excluded from this analysis, as they cannot be followed across provincial boundaries.

4.3. Diabetic Comorbidities

The present study employed the broad definition of comorbidity, which includes coexistent general medical conditions among diabetic patients and diabetic complications. To develop a new index, which measures the impact of comorbidities on hospital mortality among diabetic patients, initial lists (Appendix 3) of 48 general medical conditions and diabetic complications (Appendix 4) were selected based on review of published studies (4; 12; 13; 74) and consulting a diabetologist, Dr. R. Sigal.

4.4. Creation of a Database for the Comorbidity Index Development

All hospital separation records for each patient during the period from April 1, 1995 to March 31, 2001 were sorted chronologically and merged, based on a unique patient identifier, and sorted chronologically. Records of hospital stays for each patient were thus linked. Because transfers are considered as separate admissions in the HPOI database, some patients' admission dates may be the same as or prior to the separation dates of their previous hospital visits. Episodes of hospitalization for each patient were assembled by grouping congruent visits together. One episode of hospitalization during the study period for each patient was randomly selected for each patient. For each diabetic inpatient, all the medical records related to this episode were used to identify his/her comorbidities using ICD-9 codes. Each comorbidity was coded as a dichotomous variable (condition present or absent) to reflect whether the condition was identified in inpatient discharge abstract forms. In-hospital death was defined as death occurring for the randomly selected episode of hospitalization.

The study population of diabetic inpatients was randomly and equally split into two parts. One was used for index development and another one was used for index validation. Stratification or blocking random selection can be used for studies with small populations to ensure similar distributions in the two populations. Due to the large size of the study population in the present study, stratification or blocking by age, sex, province or calendar year was not necessary. The distributions of age, sex, province, calendar year, comorbidities and in-hospital death were similar for these two groups.

4.5. Statistical Analysis

4.5.1. Descriptive analysis

We first examined the age-specific distributions of diabetic patients with 0, 1, 2, 3, 4 and 5+ co-existent general medical conditions and diabetic complications were calculated. The percentages of in-hospital death according to the number of co-existent general medical conditions and diabetic complications (0, 1, 2, 3, 4 and 5+) were also calculated.

4.5.2. Development of the study-derived index

To develop a new index, we evaluated the contribution of each of the comorbidities in the lists of general medical conditions and diabetic complications to in-hospital mortality. Modeling was carried out using a randomly selected half of the patient population. Univariate and multiple logistic regression analyses were used to assess the contribution of each of the general medical conditions and diabetic complications to predicting in-hospital mortality in two age-groups (45-64 and 65+ years) as well as among all diabetic inpatients (45+ years). Because the differences in health care system among the provinces can influence deaths in hospital, the categorical variable "province" was included in the models. Multiple logistic regression models contained age, sex and province and comorbidities. The comorbidities significantly associated with hospital mortality were included into the new index. Following the method proposed by Charlson et al (16), we assigned the estimated odds ratios as their weights: comorbidities with $1.2 \leq$ odds ratio < 1.5 were assigned a weight of 1; comorbidities with $1.5 \leq$ odds ratio < 2.5 were assigned a weight of 2; comorbidities with $2.5 \leq$ odds ratio < 3.5 were assigned a

weight of 3;..... and comorbidities with an odds ratio > 6.5 were assigned a weight of 7. Comorbidities with adjusted odds ratio of less than 1.2 were excluded from the index. For the reasons of comparability and simplicity, the total index score for each patient was defined as the sum of weights of all comorbidities the patient experienced. For example, a diabetic inpatient aged 60 years having myocardial infarction, cerebrovascular disease and coagulopathy and severe liver disease got a study-derived score of 9 (the score = 2*(myocardial infarction) + 2*(cerebrovascular disease) + 1*(coagulopathy) + 4*(severe liver disease) = 2*1+2*1+1*1+4*1 = 9). The indices developed in different age groups were compared regarding to their predictive performance. We found that the index developed in the study population of all ages (45 years or over) performed as good as the age-specific indices, and therefore, it was selected because of its convenience of use.

4.5.3. Validation of the study-derived index

The remaining half of population not used in developing the new index (the "testing" population) was employed to assess the indices' predictive ability. We compared the study-derived index with D'Hoore-Charlson index because: 1) Of the commonly used modified versions of Charlson index, D'Hoore-Charlson index was the only one adapted to use ICD-9 diagnosis codes that was used in our data; 2) it was originally developed to predict in-hospital mortality; and 3) it assigns the original Charlson comorbidity weights to the comorbid conditions instead of the weights developed for their specific study population.

All discharge records related to the randomly selected episode of hospitalization were used to calculate the following scores for each patient: 1) the study-derived score: comorbidities that a patient experienced were assigned the study-specific weights derived

in the present study, and the sum of the weights was the study-derived score for the patient; 2) the D'Hoore-Charlson score: comorbidities in D'Hoore-Charlson index were identified and weighted by Charlson's suggested weights, and the weights of all comorbidities a patient experienced were summed up to obtain the score of D'Hoore - Charlson index for the patient; 3) the number of comorbidities: the total number of co-existent general medical conditions and diabetic complications a patient experienced was counted. The study-derived index, D'Hoore-Charlson index and the number of comorbidities were all treated as categorical variables. Since the statistic that we used to assess contribution to the model (G^2) is affected by the number of independent variables in the model, we categorized the two indices and the number of comorbidities into eleven levels (0, 1, 2, …, 10+). First, the age- and sex-specific percentages of diabetic patients with the study-derived scores, D'Hoore-Charlson scores and the number of comorbidities of 0-10+ were calculated. Second, we determined the mortality associated with increasing scores in the study-derived index (0-10+) as well as D'Hoore-Charlson Index and the number of comorbidities (0-10+) in all the age- and sex-specific groups. Finally, we compared the predictive ability of the study-derived index to that of D'Hoore-Charlson index and the number of comorbidities in all the age- and sex-specific groups.

Logistic regression models were used to examine the relative importance and the ability of the two comorbidity indices and the number of comorbidities for predicting in-hospital mortality by calculating c statistic and likelihood ratio chi-square statistic (G^2). Measures of discrimination compare the predicted outcome with the actual outcome, e.g. the c statistic, which is equivalent to the area under curve (AUC) of a receiver operating characteristic (ROC). As mentioned previously, c statistic of 0.7-0.8 is considered

acceptable and those of 0.8-0.9 excellent (75). The baseline logistic regression model was fitted to the data by modeling in-hospital death as a function of age, sex and province. For each of the two comorbidity indices and the number of comorbidities, a model was constructed containing the index, age, sex and province. The contribution of each index was assessed via G^2 resulting from comparing a model containing age, sex, province and one of the two indices or the number of comorbidities with the baseline logistic regression model that exclude the index or the number of comorbidities. When G^2 is found to be larger than expected by chance alone, we conclude that the index contributes significantly to the mortality model.

4.5.4. Comparison of the new index and the D'Hoore-Charlson index among diabetic inpatients with primary diagnosis of diabetes

We further compared the study-derived index and the D'Hoore -Charlson index in predicting in-hospital mortality and long length of stay among diabetic patients with primary diagnosis of diabetes in the "testing" population. Long length of stay was defined as length of stay at hospital for the randomly selected episode of hospitalization longer than or equal to the 75 percentile of length of stay in the age-specific groups, and coded as a dichotomous variable. The distributions of the diabetic inpatients, death in hospital and long length of stay by the scores of the two indices were calculated. The odds ratios, c statistic and likelihood ratio chi-square statistic (G^2) obtained from the multiple logistic regression models incorporating either the new index or the D'Hoore-Charlson index were compared.

5. RESULTS

5.1. Distribution of the Study Population

There were a total of 578,222 inpatients aged 45 years or over, who had either a primary or secondary diagnosis of diabetes in acute hospitals of all ten Canadian provinces, during the period from 1995/96 to 2000/01. The patients aged 45-64 years (younger group) and 65 years or over (older age group) accounted for 32.5% and 67.5%, respectively. 52.0% of the total diabetic inpatients were men, while 57.8% in the younger age group and 49.2% in the older age group were men. The overall proportion of death was 11.0% for the selected inpatients (45-64-year age group: 4.8%; 65+ year age group: 13.9%). Table 2 shows the distribution of the patients by the number of general medical conditions and diabetic complications identified in the randomly chosen episode of hospitalization. More diabetic inpatients had at least one complication in the older group than in the younger group (67.7% vs. 54.0%); a similar distribution was found for general medical conditions, which 73.1% of those aged 65 years or over and 67.8% of those aged 45-64 years had at least one general medical condition. The whole population was randomly split into half: one half for developing an index ("development" population) and the other half for validation of the index ("testing" population). The distributions of patients in the two groups were very similar in terms of the distribution of age, sex, the number of diabetic complications and the number of general medical conditions.

Table 2. Percentage (%) of diabetic patients by age, sex and number of diabetic complications and coexistent general medical conditions based on the randomly selected episode of hospitalization during the period from 1995/96 to 2000/01

	"Development" Population			"Testing" Population			Total		
	45-64 years (N = 93,657)	≥65 years (N = 195,357)	≥45 years (N = 289,014)	45-64 years (N = 94,273)	≥65 years (N = 194,935)	≥45 years (N = 289,208)	45-64 years (N = 187,930)	≥65 years (N = 390,292)	≥45 years (N = 578,222)
Sex									
Male	57.8	49.2	52.0	57.9	49.2	52.0	57.8	49.2	52.0
Number of general medical conditions									
0	32.2	26.9	28.6	32.2	26.9	28.6	32.2	26.9	28.6
1	33.0	34.3	33.9	33.0	34.1	33.7	33.0	34.2	33.8
2	19.5	21.3	20.7	19.5	21.3	20.7	19.5	21.3	20.7
3	9.1	10.4	10.0	9.1	10.5	10.1	9.1	10.5	10.0
4	3.8	4.5	4.2	3.9	4.6	4.4	3.8	4.6	4.3
5+	2.4	2.7	2.6	2.3	2.6	2.5	2.4	2.6	2.5

	"Development" Population			"Testing" Population			Total		
	45-64 years (N = 93,657)	≥65 years (N = 195,357)	≥45 years (N = 289,014)	45-64 years (N = 94,273)	≥65 years (N = 194,935)	≥45 years (N = 289,208)	45-64 years (N = 187,930)	≥65 years (N = 390,292)	≥45 years (N = 578,222)
Number of complications									
0	46.0	32.3	36.8	46.0	32.2	36.7	46.0	32.3	36.7
1	28.7	29.9	29.5	29.0	29.9	29.6	28.9	29.9	29.6
2	14.8	19.8	18.1	14.7	19.9	18.2	14.8	19.8	18.2
3	6.3	10.6	9.2	6.2	10.6	9.2	6.3	10.6	9.2
4	2.6	4.7	4.0	2.6	4.7	4.0	2.6	4.7	4.0
5+	1.6	2.7	2.4	1.5	2.7	2.3	1.5	2.7	2.3
Death in hospital	4.9	13.9	10.9	4.8	14.0	11.0	4.8	13.9	11.0

5.2. Development of the Study-derived Index

5.2.1. Prevalence of comorbidities in the "development" population

Table 3 shows the prevalence of co-existent general medical conditions and diabetic complications in the "development" population. These conditions were positively associated with in-hospital mortality among diabetic inpatients. The five most prevalent general medical conditions were cancer/ malignancy including leukemia and lymphoma, fluid and electrolyte disorders, mild to severe liver disease, respiratory failure/ insufficiency/ arrest, and septicemia in both younger and older age groups. Liver disease and acquired immune deficiency syndrome were more prevalent in the younger age group than in the older age group, while cancer/malignancy including leukemia and lymphoma, fluid and electrolyte disorders, respiratory failure/ insufficiency/ arrest, septicemia, fracture of neck of femur (hip) and lower respiratory disease were more prevalent in the older age group. The five most prevalent diabetic complications were cardiomegaly/heart failure, conduction disorders/ cardiac dysrhythmias, myocardial infarction, cerebrovascular disease and peripheral vascular disease in both age groups. The prevalences of diabetic complications were higher in the older age group, except the prevalences of neuropathies and thrombotic disease were similar in the two groups. The differences between the two age groups were greater for cardiomegaly/heart failure, conduction disorders/ cardiac dysrhythmias, and cerebrovascular disease, which were 11.5%, 10.0% and 7.3%, respectively.

Table 3. Number (%) of diabetic inpatients with complications and coexistent general medical conditions by age based on the randomly selected episode of hospitalization in the randomly selected half of population ("development" population) during the period from 1995/96 to 2000/01

	45-64 years N = 93,657	≥ 65 years N = 195,357	≥ 45 years N = 289,014
	n (%)	n (%)	n (%)
General medical conditions			
Cancer/ malignancy (incl. Leukemia & lymphoma)	8689 (9.3)	25622 (13.1)	34311 (11.9)
Fluid and electrolyte disorders	4184 (4.5)	14088 (7.2)	18272 (6.3)
Mild liver disease	2988 (3.2)	4139 (2.1)	7127 (2.5)
Moderate or severe liver disease	1297 (1.4)	1687 (0.9)	2984 (1.0)
Respiratory failure, insufficiency, arrest (adult)	1783 (1.9)	5853 (3.0)	7636 (2.6)
Septicemia (excluding during labor)	1696 (1.8)	4959 (2.5)	6655 (2.3)
Fracture of neck of femur (hip)	330 (0.4)	4511 (2.3)	4841 (1.7)
Pulmonary circulation disorders	1050 (1.1)	3010 (1.5)	4060 (1.4)
Hereditary and degenerative nervous system disorder (excluding Parkinson's and multiple sclerosis)	1316 (1.4)	3207 (1.6)	4523 (1.6)
Coagulopathy	1289 (1.4)	2716 (1.4)	4005 (1.4)
Lower respiratory disease (excluding asthma, pleurisy, and respiratory failure)	676 (0.7)	2576 (1.3)	3252 (1.1)
Epilepsy	643 (0.7)	1021 (0.5)	1664 (0.6)
Peritonitis and intestinal abscess	408 (0.4)	792 (0.4)	1200 (0.4)
Acquired immune deficiency syndrome (AIDS)	92 (0.1)	13 (0.01)	105 (0.04)

	45-64 years N = 93,657	≥ 65 years N = 195,357	≥ 45 years N = 289,014
	% (no.)	% (no.)	% (no.)
Diabetic complications			
Cardiomegaly and heart failure (group 8)	8414 9 (9.0)	39987 (20.5)	48401 (16.7)
Conduction disorders and cardiac dysrhythmias (group 9)	8028 (8.6)	36254 (18.6)	44282 (15.3)
Myocardial infarction (group7a)	12371 (13.2)	29336 (15.0)	41707 (14.4)
Cerebrovascular disease (group 2)	6210 (6.6)	27058 (13.9)	33268 (11.5)
Peripheral vascular disease (group 5)	7923 (8.5)	20628 (10.6)	28551 (9.9)
Renal failure (group 10b)	4177 (4.5)	16715 (8.6)	20892 (7.2)
Neuropathies (group 1)	3119 (3.3)	6382 (3.3)	9501 (3.3)
Thrombotic disease (group 4)	1814 (1.9)	4180 (2.1)	5994 (2.1)

5.2.2. Distribution of death in hospital among the "development" population

Table 4 presents the in-hospital mortality by the number of co-existent general medical conditions and diabetic complications in age-specific groups in the "development" population. The overall in-hospital mortality was 10.9%, compared with 4.8% in the younger age group and 13.9% in the older age group. In both age groups, the hospital mortality was higher in those with comorbid condition(s) compared with those without. The in-hospital mortality increased steadily with increasing number of diabetic complications or number of coexistent general medical conditions. The hospital mortality were 3- and 6-fold higher among diabetic patients with 5 or more co-existent general medical conditions than those with only one for the younger and older age groups, respectively. The corresponding crude relative risk estimates were 2.0 and 3.6 for those with 5 or more diabetic complications than those with only one complication, respectively.

Table 4. Number (%) of deaths by age, number of complications and coexistent general medical conditions among diabetic inpatients based on the randomly selected episode of hospitalization in the randomly selected half of population ("development" population) during the period from 1995/96 to 2000/01

	45-64 years		≥ 65 years		≥ 45 years	
	N	n. (%)	N	n. (%)	N	n. (%)
Number of general medical conditions						
0	43045	1318 (3.1)	63245	4986 (7.9)	106290	6304 (5.9)
1	26918	1027 (3.8)	58376	6540 (11.2)	85294	7567 (8.9)
2	13919	849 (6.1)	38508	6363 (16.5)	52427	7212 (13.8)
3	5881	630 (10.7)	20823	4807 (23.1)	26704	5437 (20.4)
4	2416	378 (15.6)	9095	2574 (28.3)	11511	2952 (25.6)
5+	1478	336 (22.7)	5310	1822 (34.3)	6788	2158 (31.8)
Total	93657	4538 (4.8)	195357	27092 (13.9)	289014	31630 (10.9)
Number of complications						
0	30199	636 (2.1)	52456	5346 (10.2)	82655	5982 (7.2)
1	30911	1345 (4.4)	67084	8605 (12.8)	97995	9950 (10.2)
2	18226	1079 (5.9)	41541	6385 (15.4)	59767	7464 (12.5)
3	8533	719 (8.4)	20280	3644 (18.0)	28813	4363 (15.1)
4	3508	402 (11.5)	8828	1763 (20.0)	12336	2165 (17.6)
5+	2280	357 (15.7)	5168	1349 (26.1)	7448	1706 (22.9)
Total	93657	4538 (4.8)	195357	27092 (13.9)	289014	31630 (10.9)

5.2.3. The weight for each variable in the study-derived Index

Table 5 shows the odds ratios for hospital mortality associated with comorbidity (including co-existent general medical conditions and diabetic complications) among diabetic inpatients in the "development" population. Fourteen out of 48 coexistent general medical conditions and 8 out of 14 diabetic complications had significantly positive association with death in hospital among diabetic inpatients aged 45 years or over. Respiratory failure, peritonitis and intestinal abscess, and cancer/ malignancy including leukemia and lymphoma had a strong association, followed by moderate or severe liver disease, septicemia, acquired immune deficiency syndrome and renal failure.

Some comorbidities associated with hospital mortality were age related. Fracture of neck of the femur (hip) and neuropathies were associated with 46% and 52% of increase in the odds of hospital death in the older age group, but no significant associations were observed in the younger age group. Hereditary and degenerative nervous system disorder, epilepsy and peripheral vascular disease were significantly associated with an increased odds of death in hospital by 115%, 101% and 62% respectively in the younger age group, and no significant impact was observed in the older age group. For all the comorbid conditions significantly associated with hospital mortality in both age groups, eight of them might have a stronger association in the younger age group, particularly cancer/malignancy including leukemia and lymphoma. Five small hospitals in Prince Edward Island and all hospitals in Saskatchewan only report primary diagnosis and 2-3 secondary diagnoses. When we excluded the patients from these two provinces and repeated the analyses, the odds ratios and the predictive performance of the model measured by c statistic had no notable changes.

The study-derived index was developed by including the coexistent general medical conditions and diabetic complications, which were significantly associated with an increased odds of death in hospital. Estimated odds ratios were assigned as weights for the index. The odds ratio was rounded to the nearest integer. (See Table 6 for the study-derived index.) We compared the study-derived index developed for diabetic inpatients aged 45 years or over to the index developed specifically for those either aged 45-64 years or 65 years or over of "testing" population, respectively, with the c statistic, a measure of discrimination, being almost the same.

Table 5. Adjusted odds ratios and the 95% confidence intervals of hospital mortality for comorbidities and the index derived from the odds ratios based on the randomly selected episode of hospitalization in the randomly selected half of population ("development" population) during the period from 1995/96 to 2000/01

	45-64 years		≥65 years		≥45 years	
	Adjusted OR	Assigned weight	Adjusted OR	Assigned weight	Adjusted OR	New index (assigned weight)
Cancer/ malignancy (incl. leukemia & lymphoma)	11.36 (10.49, 12.29)	7	4.36 (4.20, 4.52)	4	5.15 (4.99, 5.32)	5
Fluid and electrolyte disorders	1.63 (1.45, 1.84)	2	1.44 (1.37, 1.51)	1	1.47 (1.40, 1.53)	1
Mild liver disease	2.08 (1.79, 2.43)	2	1.55 (1.41, 1.70)	1	1.61 (1.49, 1.74)	2
Moderate or sever liver disease	3.99 (3.30, 4.81)	4	4.30 (3.79, 4.87)	4	4.16 (3.75, 4.61)	4
Respiratory failure, insufficiency, arrest (adult)	5.19 (4.55, 5.92)	5	5.94 (5.59, 6.32)	6	5.84 (5.52, 6.16)	6
Septicemia (excluding during labor)	3.92 (3.39, 4.53)	4	3.83 (3.58, 4.10)	4	3.86 (3.63, 4.11)	4

	45-64 years		≥65 years		≥45 years	
	Adjusted OR	**Assigned weight**	**Adjusted OR**	**Assigned weight**	**Adjusted OR**	**New index (assigned weight)**
Fracture of neck of femur (hip)	-	-	1.46 (1.34, 1.59)	1	1.52 (1.39, 1.65)	1
Pulmonary circulation disorders	3.14 (2.62, 3.78)	3	2.06 (1.88, 2.25)	2	2.24 (2.06, 2.43)	2
Hereditary and degenerative nervous system disorder (excluding Parkinson's and multiple sclerosis)	2.15 (1.75, 2.63)	2	-	-	1.24 (1.13, 1.36)	1
Coagulopathy	1.87 (1.56, 2.24)	2	1.37 (1.23, 1.51)	1	1.48 (1.36, 1.62)	1
Lower respiratory disease (excluding asthma, pleurisy, and respiratory failure)	3.23 (2.58, 4.05)	3	2.08 (1.89, 2.29)	2	2.25 (2.06, 2.46)	2
Epilepsy	2.01 (1.48, 2.73)	2	-	-	1.33 (1.14, 1.56)	1
Peritonitis and intestinal abscess	3.28 (2.44, 4.39)	3	4.26 (3.61, 5.03)	4	3.96 (3.43, 4.57)	4
Acquired immune deficiency syndrome (AIDS)	6.56 (3.57, 12.06)	7	4.41 (1.23, 15.81	4	6.01 (3.54, 10.22)	6

	45-64 years		≥65 years		≥45 years	
	Adjusted OR	**Assigned weight**	**Adjusted OR**	**Assigned weight**	**Adjusted OR**	**New index (assigned weight)**
Neuropathies (group 1)	-	-	1.52 (1.42, 1.64)	1	1.32 (1.23, 1.42)	1
Cerebrovascular disease (group 2)	3.18 (2.87, 3.53)	3	2.39 (2.30, 2.47)	2	2.45 (2.37, 2.54)	2
Thrombotic disease (group 4)	1.46 (1.21, 1.77)	1	1.36 (1.25, 1.49)	1	1.37 (1.26, 1.49)	1
Peripheral vascular disease (group 5)	1.62 (1.46, 1.80)	2	-	-	1.26 (1.21, 1.32)	1
Myocardial infarction (group7a)	2.56 (2.34, 2.79)	3	2.27 (2.20, 2.36)	2	2.30 (2.23, 2.38)	2
Cardiomegaly and heart failure (group 8)	1.75 (1.59, 1.92)	2	1.77 (1.71, 1.82)	2	1.77 (1.72, 1.83)	2
Conduction disorders and cardiac dysrhythmias (group 9)	3.69 (3.39, 4.02)	4	1.65 (1.59, 1.70)	2	1.81 (1.76, 1.87)	2
Renal failure (group 10b)	3.57 (3.22, 3.96)	4	2.71 (2.60, 2.82)	3	2.79 (2.69, 2.90)	3

Table 6. The study-derived index

Weight	Comorbidity	ICD-9 code
1	Peripheral vascular disease (group 5)	440, 443, 459, 785.4, 885–887, 895–897, 250.7
	Thrombotic disease (group 4)	444, 447.1, 451-453
	Neuropathies (group 1)	356.8, 354, 355, 337.1, 357.2, 729.2, 250.6
	Coagulopathy	286, 287.1, 287.3-287.5,
	Epilepsy	345.9
	Hereditary and degenerative nervous system disorder (excluding Parkinson's and multiple sclerosis)	330, 331.3, 331.4, 331.7-331.9, 333.0-333.8, 334, 335, 336, 337.0, 337.1, 337.9
	Fracture of neck of femur (hip)	820, 905.3
	Fluid and electrolyte disorders	276
2	Mild liver disease	571, 573
	Pulmonary circulation disorders	415, 416, 417
	Lower respiratory disease (excluding asthma, pleurisy, and respiratory failure)	513-516
	Cerebrovascular disease (group 2)	433-438
	Myocardial infarction (group7a)	410, 412
	Cardiomegaly and heart failure (group 8)	429.3, 425, 428, 429.1
	Conduction disorders and cardiac dysrhythmias (group 9)	426–427
3	Renal failure (group 10b)	584-586, 588

Weight	Comorbidity	ICD-9 code
4	Moderate or sever liver disease	070, 570, 572
	Septicemia (excluding during labor)	003.1, 020.0, 022.3, 036.2, 038, 054.5, 790.7
	Peritonitis and intestinal abscess	567, 569.5
5	Cancer/ malignancy (incl. leukemia & lymphoma)	140-195, 196-198, 199, 200-208, 230-234, 235-238, 239
6	Respiratory failure, insufficiency, arrest (adult)	518.5, 518.8, 799.1, V46.1
6	Acquired immune deficiency syndrome (AIDS)	042, 043, 044

5.3. Validation of the Study-derived Index

5.3.1. Distribution of "testing" population

Table 7 presents the distribution of diabetic inpatients for the "testing" population by the scores of the study-derived index and D'Hoore-Charlson index and the number of comorbidities in the age- and sex-specific groups. Among the diabetic patients 45 years of age or over, 64.1% and 58.1% had the study-derived score and D'Hoore-Charlson score of one or over respectively, and the proportions were lower in the younger age group than in the older age group. Overall 88.6% of diabetic patients had at least one comorbid condition, and the percentages were lower in the younger age group than in the older age group. More men than women had the study-derived score and D'Hoore-Charlson score of one or over or at least one comorbid condition, and the differences were more evident in the younger age group than in the older age group.

Table 7. Percentage (%) of diabetic patients by age, sex and index values or the number of comorbidities in the "testing" population* during the period from 1995/96 to 2000/01

Index score or number count	45-64 years			≥ 65 years			≥ 45 years		
	Male (N = 54,576) %	Female (N = 39,697) %	Total (N = 94,273) %	Male (N = 95,990) %	Female (N = 98,945) %	Total (N = 194,935) %	Male (N = 150,566) %	Female (N = 138,642) %	Total (N = 289,208) %
The study-derived Index**									
0	44.3	54.2	48.5	27.5	31.9	29.8	33.6	38.3	35.9
1	6.6	6.9	6.7	5.6	7.4	6.5	6.0	7.3	6.6
2	20.9	14.8	18.3	21.0	20.4	20.7	21.0	18.8	19.9
3	4.0	3.2	3.7	5.3	5.5	5.4	4.8	4.8	4.8
4	6.6	4.6	5.8	9.8	9.5	9.6	8.6	8.1	8.4
5	7.5	8.5	7.9	11.1	9.6	10.3	9.8	9.3	9.5
6	3.3	2.6	3.0	5.3	4.6	5.0	4.6	4.1	4.3
7	2.1	1.6	1.9	4.8	3.9	4.3	3.8	3.2	3.5
8	1.2	1.1	1.2	2.7	2.1	2.4	2.1	1.8	2.0
9	1.0	0.8	0.9	2.3	1.7	2.0	1.8	1.4	1.6
10+	2.4	1.8	2.2	4.8	3.5	4.1	3.9	3.0	3.5
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Index score or number count	45-64 years			≥ 65 years			≥ 45 years		
	Male (N = 54,576) %	Female (N = 39,697) %	Total (N = 94,273) %	Male (N = 95,990) %	Female (N = 98,945) %	Total (N = 194,935) %	Male (N = 150,566) %	Female (N = 138,642) %	Total (N = 289,208) %
D'Hoore-Charlson Index									
0	50.0	56.7	52.9	34.2	39.1	36.7	39.9	44.1	41.9
1	24.7	20.6	23.0	24.8	25.3	25.1	24.8	24.0	24.4
2	13.6	12.5	13.1	19.4	18.1	18.7	17.3	16.5	16.9
3	4.6	3.9	4.3	9.2	8.0	8.6	7.5	6.8	7.2
4	2.4	1.6	2.1	4.6	3.6	4.1	3.8	3.0	3.4
5	0.7	0.5	0.6	1.7	1.2	1.5	1.3	1.0	1.2
6	0.9	1.0	1.0	1.4	1.2	1.3	1.2	1.1	1.2
7	0.2	0.2	0.2	0.4	0.3	0.4	0.3	0.3	0.3
8	2.1	2.5	2.3	2.9	2.3	2.6	2.6	2.3	2.5
9	0.4	0.3	0.4	0.7	0.5	0.6	0.6	0.5	0.5
10+	0.3	0.2	0.3	0.8	0.5	0.6	0.6	0.4	0.5
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Index score or number count	45-64 years			≥ 65 years			≥ 45 years		
	Male (N = 54,576) %	Female (N = 39,697) %	Total (N = 94,273) %	Male (N = 95,990) %	Female (N = 98,945) %	Total (N = 194,935) %	Male (N = 150,566) %	Female (N = 138,642) %	Total (N = 289,208) %
Number of Comorbidities									
0	15.0	17.7	16.1	9.4	8.9	9.2	11.4	11.4	11.4
1	24.8	25.2	25.0	21.0	19.9	20.4	22.4	21.5	21.9
2	23.3	22.4	22.9	22.8	22.4	22.6	23.0	22.4	22.7
3	15.9	14.9	15.5	17.9	18.4	18.1	17.2	17.4	17.3
4	9.5	8.7	9.2	12.0	12.5	12.2	11.1	11.4	11.2
5	5.5	5.4	5.5	7.7	8.0	7.8	6.9	7.3	7.1
6	3.2	3.0	3.1	4.8	4.9	4.8	4.2	4.3	4.3
7	1.7	1.5	1.6	2.6	2.7	2.6	2.3	2.3	2.3
8	0.8	0.8	0.8	1.3	1.4	1.4	1.1	1.2	1.2
9	0.3	0.4	0.3	0.5	0.6	0.6	0.5	0.5	0.5
10+	0.1	0.1	0.1	0.2	0.2	0.2	0.2	0.2	0.2
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

* The Testing population consists of a randomly selected episode of hospitalization for each patient in the randomly selected half of the population.

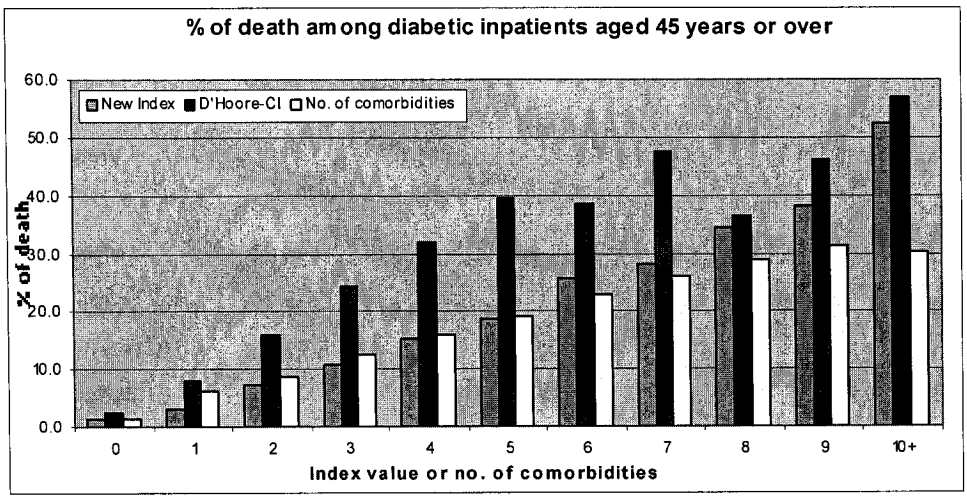
** The study-derived index refers to the index derived from diabetic inpatients aged 45 years or over in the present study.

5.3.2. Distribution of death in hospital among the "testing" population

Figures 1-9 show hospital mortality by the scores of the study-derived index and the D'Hoore-Charlson index and the number of comorbidities in age- and sex-specific groups. Diabetic patients with a higher score of the study-derived index had higher in-hospital mortality in all age- and sex-specific groups. Diabetic men and women had similar hospital mortality across the age groups. The linear pattern was more consistent for the study-derived index compared with D'Hoore-Charlson score or the number of comorbidities in all the age and sex-specific groups. Diabetic patients with a score of the D'Hoore-Charlson index of 6 or 8 had lower in-hospital mortality than those with a score of 5.

Figure 1-9. Crude percentage of death by index values or the number of comorbidities in the "testing" population* during the period from 1995/96 to 2000/01

Figure 1.



* The Testing population consists of a randomly selected episode of hospitalization for each patient in the randomly selected half of the population.

Figure 2.

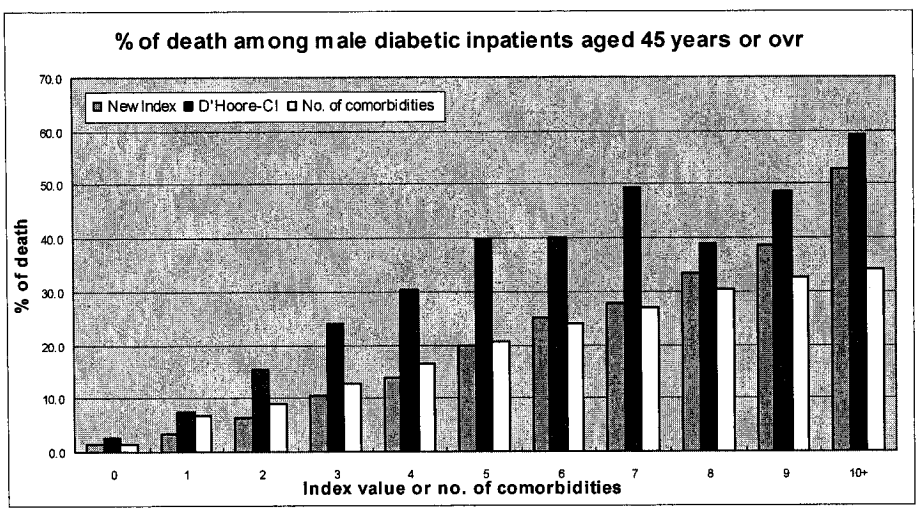


Figure 3.

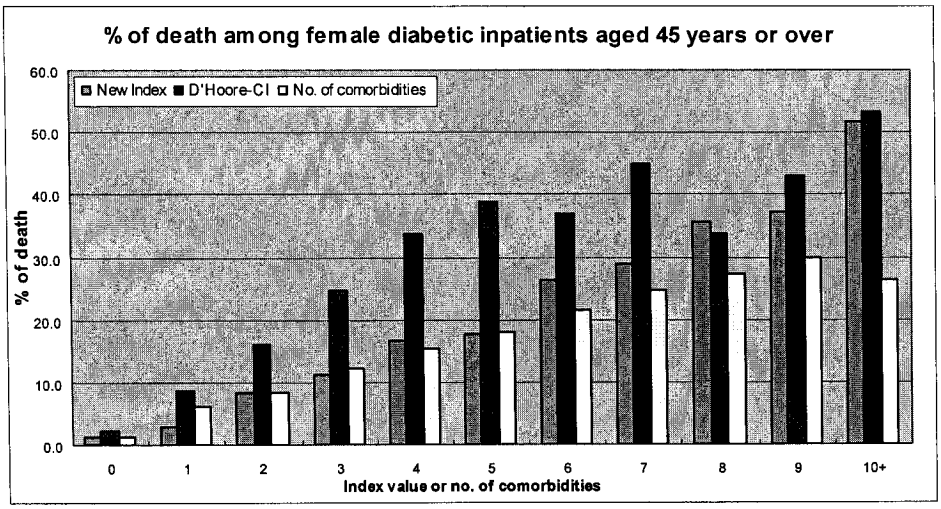


Figure 4.

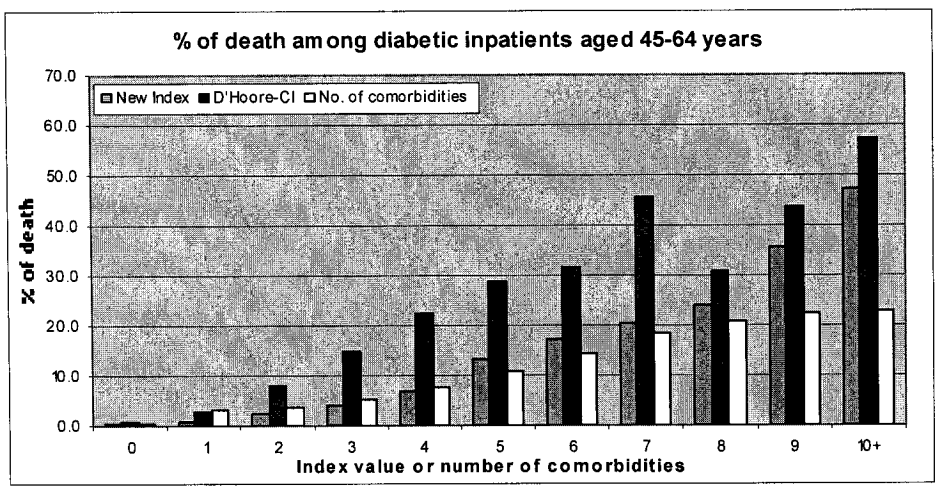


Figure 5.

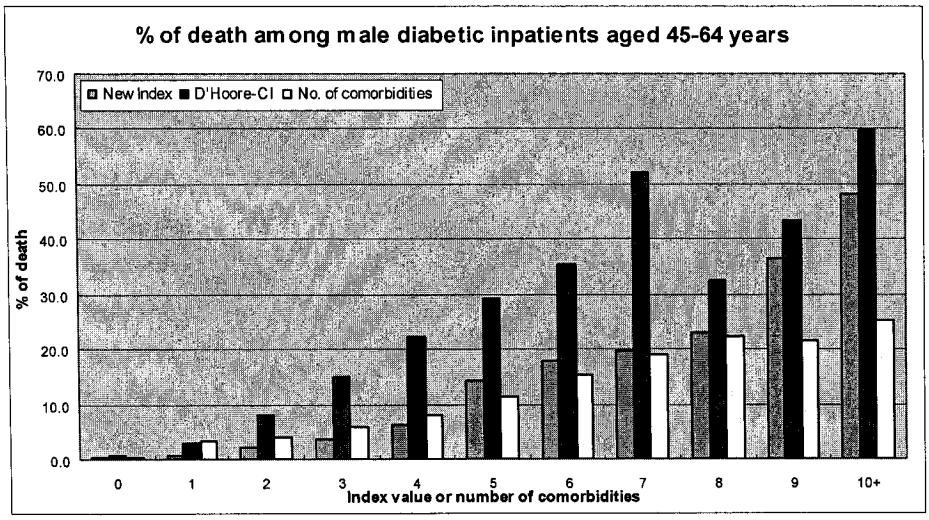


Figure 6.

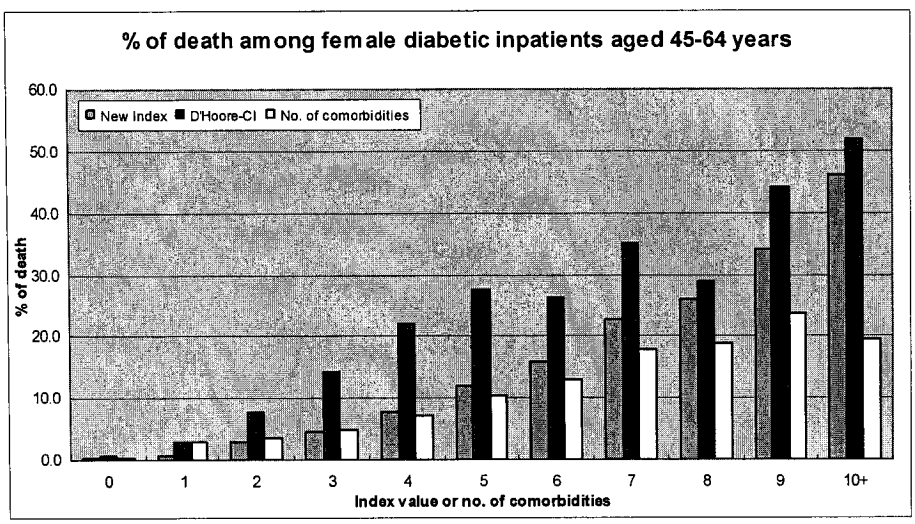


Figure 7.

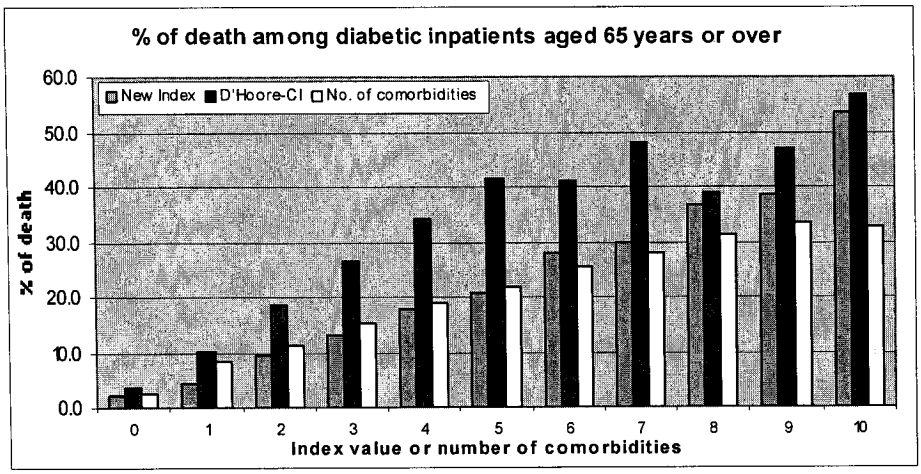


Figure 8.

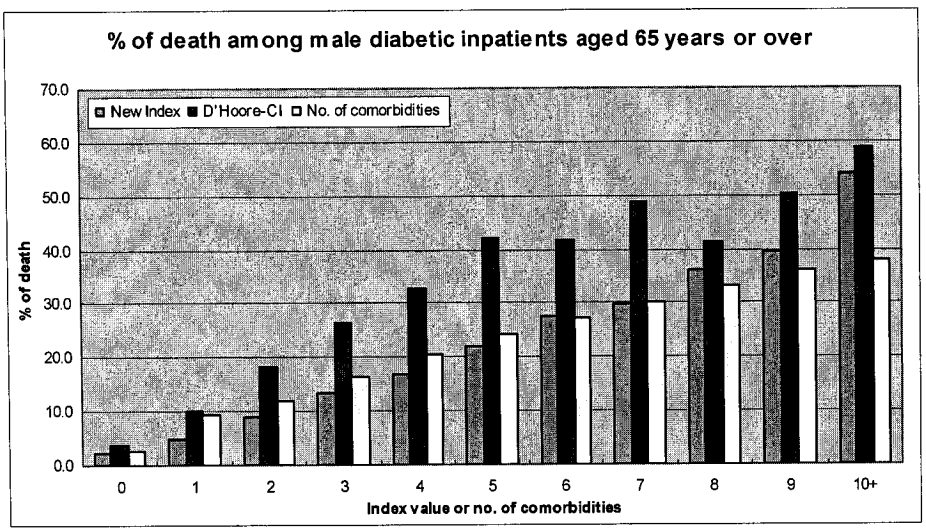
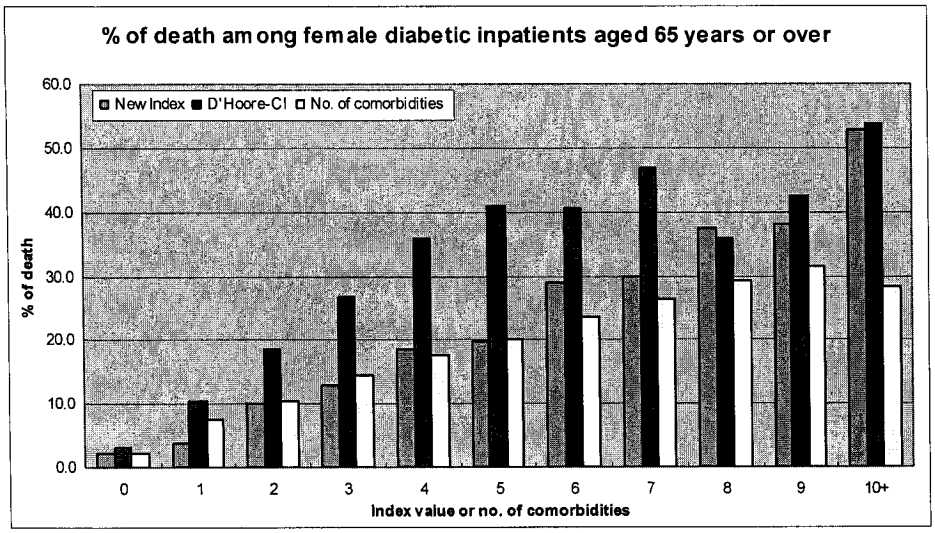


Figure 9.


% of death among female diabetic inpatients aged 65 years or over
New Index
D'Hoore-CI
No. of comorbidities
60.0
50.0
40.0
30.0
20.0
10.0
0.0
% of death
0
1
2
3
4
5
6
7
8
9
10+
Index value or no. of comorbidities

5.3.3. Associations between the indices and in-hospital mortality

Table 8 shows the odds ratios for hospital mortality associated with the scores of the study-derived index, D'Hoore-Charlson index and the number of comorbidities among diabetic inpatients in the "testing" population. The scores and the number of comorbidities were treated as categorical variables. In all the age groups, the scores were significantly associated with death in hospital. The higher the scores or the number of comorbidities were, the higher the odds ratios were. The odds of death in hospital were two, nearly five and seven times higher among diabetic inpatients aged 45 years or over who scored 1, 2 or 3 on the study-derived index than those who had no significant comorbid conditions after controlling for age, sex and province. For the two indices and the number of comorbidities, the odds ratios were significantly higher among middle-aged diabetic inpatients than their older counterparts with the same scores or the number of comorbidities; and the odds ratios were not significantly different between diabetic men and women in all the age groups (see appendix 4). The odds of death in hospital tended to increase more markedly with increasing scores of the indices and number of comorbidities in the younger age group than in the older age group.

Table 8. Adjusted odds ratios and the 95% confidence intervals for hospital mortality associated with the indices in the "testing" population* during the period from 1995/96 to 2000/01

	The study-derived Index**	D'Hoore-Charlson Index	Number of Comorbidities
45-64 years			
0	1.00	1.00	1.00
1	1.97 (1.45, 2.69)	3.84 (3.38, 4.36)	7.39 (5.70, 9.59)
2	5.91 (4.97, 7.03)	10.77 (9.55, 12.14)	9.17 (7.08, 11.89)
3	10.17 (8.14, 12.70)	21.66 (18.96, 24.75)	13.60 (10.49, 17.64)
4	16.74 (14.00, 20.03)	35.92 (30.98, 41.64)	20.07 (15.44, 26.09)
5	35.76 (30.51, 41.91)	50.40 (40.99, 61.96)	30.25 (23.20, 39.45)
6	49.40 (41.51, 58.80)	55.40 (46.57, 65.91)	41.48 (31.62, 54.42)
7	60.59 (50.39, 72.86)	104.63 (74.99, 145.99)	56.72 (42.73, 75.28)
8	75.26 (61.56, 92.01)	55.36 (48.28, 63.48)	67.47 (49.49, 91.97)
9	130.19 (106.35, 159.38)	97.99 (77.00, 124.70)	76.04 (52.63, 109.86)
10+	216.25 (182.67, 256.00)	169.15 (129.97, 220.14)	75.29 (46.20, 122.69)
≥ 65 years			
0	1.00	1.00	1.00
1	1.89 (1.68, 2.06)	3.02 (2.88, 3.18)	3.82 (3.45, 4.23)
2	4.38 (4.10, 4.67)	6.02 (5.73, 6.31)	5.27 (4.77, 5.83)
3	6.39 (5.90, 6.92)	9.65 (9.15, 10.18)	7.76 (7.01, 8.58)
4	8.98 (8.39, 9.60)	14.44 (13.57, 15.37)	10.30 (9.30, 11.41)
5	11.77 (11.03, 12.57)	20.25 (18.57, 22.08)	12.73 (11.47, 14.13)

	The study-derived Index**	D'Hoore-Charlson Index	Number of Comorbidities
6	17.23 (16.04, 18.51)	20.12 (18.28, 21.92)	15.90 (14.27, 17.72)
7	18.54 (17.23, 19.95)	26.46 (22.59, 30.99)	19.04 (16.97, 21.37)
8	26.24 (24.16, 28.50)	19.66 (18.32, 21.09)	23.26 (20.45, 26.46)
9	28.15 (25.81, 30.70)	26.78 (23.71, 30.25)	26.18 (22.29, 30.76)
10+	54.45 (50.67, 58.50)	40.39 (35.72, 45.67)	26.87 (21.13, 34.17)
≥ 45 years			
0	1.00	1.00	1.00
1	1.97 (1.79, 2.17)	3.18 (3.04, 3.33)	4.29 (3.90, 4.72)
2	4.74 (4.46, 5.04)	6.64 (6.35, 6.94)	5.83 (5.31, 6.40)
3	7.13 (6.62, 7.69)	10.98 (10.45, 11.53)	8.60 (7.83, 9.45)
4	10.19 (9.57, 10.86)	16.74 (15.81, 17.73)	11.61 (10.56, 12.77)
5	14.42 (13.58, 15.32)	23.43 (21.62, 25.38)	14.79 (13.42, 16.30)
6	20.79 (19.46, 22.21)	24.38 (22.50, 26.42)	18.78 (16.99, 20.76)
7	22.60 (21.11, 24.20)	33.54 (29.03, 38.76)	23.03 (20.70, 25.62)
8	31.51 (29.19, 34.01)	24.35 (22.89, 25.91)	28.02 (24.88, 31.56)
9	36.65 (33.83, 39.70)	34.14 (30.60, 38.09)	31.67 (27.32, 36.71)
10+	70.59 (66.08, 75.40)	52.10 (46.56, 58.30)	32.65 (26.31, 40.52)

* The testing population consists of a randomly selected episode of hospitalization for each patient in the randomly selected half of the population

** The study-derived index refers to the index derived from diabetic inpatients aged 45 years or over in the present study.

5.3.4. Predictive performance

Table 9 shows the c statistic and the likelihood ratio chi-square statistic (G^2) of the models. Because the number of categories of a categorical variable influences the c statistic and G^2, the two indices and the number of comorbidities were truncated at the value of 10 for a fair comparison. The eleven-level comorbidity index or comorbidity number was transformed into ten dummy variables. Among diabetic inpatients aged 45 years or over, the study-derived index performed best in predicting death in hospital, and the c statistic was 0.83 for the multivariate model incorporating the study-derived index. This finding represents an improvement of 0.16 (23.9 %) over the baseline model alone including only age, sex and province (c = 0.67). The c statistics were lower for the model incorporating the D'Hoore-Charlson Index (c = 0.81), and much lower for the model incorporating the number of comorbidities (c = 0.74). As compared with the baseline regression model, the addition of the study-derived index resulted in a significant improvement in the fit (ΔG^2 = 189789.9-155762.2 = 34027.7, 10 d.f., $p < 0.001$). Adding the D'Hoore-Charlson index or the number of comorbidities also improved the fit of the model (D'Hoore-Charlson index: ΔG^2 = 189789.9-162079.6 = 27710.3, 10 d.f., $p < 0.001$; as did adding the number of comorbidities: ΔG^2 = 189789.9-178647.1 = 11142.8, 10 d.f., $p < 0.001$). The deviance for the study-derived index was lower than that for the D'Hoore-Charlson index (155762.2 versus 162079.6), indicating a better predictive ability for the study-derived index. The study-derived index, D'Hoore-Charlson index and the number of comorbidities all performed better among the younger diabetic inpatients than among older diabetic inpatients, and their predictive abilities were very similar for diabetic men and women in all age groups.

Table 9: Comparison of the study-derived index*, D'Hoore-Charlson index and the number of comorbidities in predicting death in hospital using the "testing" population during the period from 1995/96 to 2000/01**

Models	45-64 years		≥65 years		≥ 45 years	
	C statistic	G^2 (-2 Log L / d.f.)	C statistic	G^2 (-2 Log L / d.f.)	C statistic	G^2 (-2 Log L / d.f.)
MALE + FEMALE						
Baseline model (Age, sex, province)	0.61	35845.9 / 11	0.61	153883.8 / 11	0.67	189789.9 / 11
Baseline model + the study-derived index*	0.88	26065.6 / 21	0.80	129031.5 / 21	0.83	155762.2 / 21
Baseline model + D'Hoore-CI	0.85	28008.9 / 21	0.78	133595.4 / 21	0.81	162079.6 / 21
Baseline model + number of comorbidities	0.74	33024.1 / 21	0.70	145457.2 / 21	0.74	178647.1 / 21
MALE						
Baseline model (Age, sex, province)	0.61	21840.9 / 10	0.60	79092.9 / 10	0.67	100980.8 / 10
Baseline model + the study-derived index*	0.88	15917.6 / 20	0.80	66480.6 / 20	0.83	82812.9 / 20
Baseline model + D'Hoore-CI	0.85	17058.7 / 20	0.77	68782.8 / 20	0.81	86144.0 / 20
Baseline model + number of comorbidities	0.73	20180.5 / 20	0.70	74557.8 / 20	0.74	94819.6 / 20
FEMALE						
Baseline model (Age, sex, province)	0.60	13995.1 / 10	0.62	74776.6 / 10	0.67	88787.1 / 10
Baseline model + the study-derived index*	0.89	10112.1 / 20	0.80	62496.0 / 20	0.83	72875.1 / 20
Baseline model + D'Hoore-CI	0.85	10930.2 / 20	0.78	64760.8 / 20	0.81	75887.7 / 20
Baseline model + number of comorbidities	0.74	12826.5 / 20	0.70	70868.7 / 20	0.74	83781.5 / 20

* The study-derived index refers to the index derived from diabetic inpatients aged 45 years or over in the present study.

** The testing population consists of a randomly selected episode of hospitalization for each patient in the randomly selected half of the population

5.4. Comparison of the Study-derived Index and D'Hoore-Charlson Index among Diabetic Patients with Primary Diagnosis of Diabetes

5.4.1. Distribution of diabetic patients with primary diagnosis of diabetes in the "testing" population

Table 10 presents the distribution of diabetic inpatients with primary diagnosis of diabetes in the "testing" population by the scores of the study-derived index and D'Hoore-Charlson index in the age-specific groups. Of 289,208 diabetic inpatients in the "testing" population, 22,821 (7.9%) patients were admitted to hospital due to diabetes as primary diagnosis (ICD-9 code: 250). The patients aged 45-64 years accounted for 43.4% while the patients aged 65 years or over accounted for 56.6%. Among the diabetic patients 45 years of age or over, 55.2% and 36.7% had the study-derived score and D'Hoore-Charlson score of one or over respectively, and the proportions were lower in the younger age group than in the older age group.

Table 10. Percentage (%) of patients with primary diagnosis of diabetes by index values in the "testing" population* during the period from 1995/96 to 2000/01

Index score	45-64 years (N = 9,903)		≥ 65 years (N = 12,918)		≥ 45 years (N = 22,821)	
	The study-derived Index** (%)	D'Hoore-Charlson Index (%)	The study-derived Index (%)	D'Hoore-Charlson Index (%)	The study-derived Index (%)	D'Hoore-Charlson Index (%)
0	54.7	72.6	37.3	56.2	44.8	63.3
1	16.0	11.5	15.6	19.5	15.7	16.0
2	12.6	9.6	17.6	13.0	15.5	11.5
3	4.7	3.6	7.5	6.0	6.3	5.0
4	3.4	1.3	6.2	2.8	5.0	2.1
5	3.4	0.3	5.3	0.9	4.5	0.6
6	2.0	0.4	3.4	0.5	2.8	0.4
7	1.2	0.1	2.5	0.1	1.9	0.1
8	0.7	0.5	1.6	0.8	1.2	0.6
9	0.5	0.1	1.1	0.2	0.8	0.1
10+	1.0	0.1	1.9	0.1	1.5	0.1
Total	100	100	100	100	100	100

* The testing population consists of a randomly selected episode of hospitalization for each patient in the randomly selected half of the population.

** The study-derived index refers to the index derived from diabetic inpatients aged 45 years or over in the present study.

5.4.2. Distribution of death in hospital among diabetic patients with primary diagnosis of diabetes in the "testing" population

Figures 10-12 show the age-specific hospital mortality by scores of the study-derived index and the D'Hoore-Charlson index. The diabetic patients with a higher score of the study-derived index had higher in-hospital mortality in all age groups. This linear pattern was not evident for the D'Hoore-Charlson score.

Figure 10-12: Crude percentage of death among patients with primary diagnosis of diabetes by index values in the "testing" population* during the period from 1995/96 to 2000/01

Figure 10.

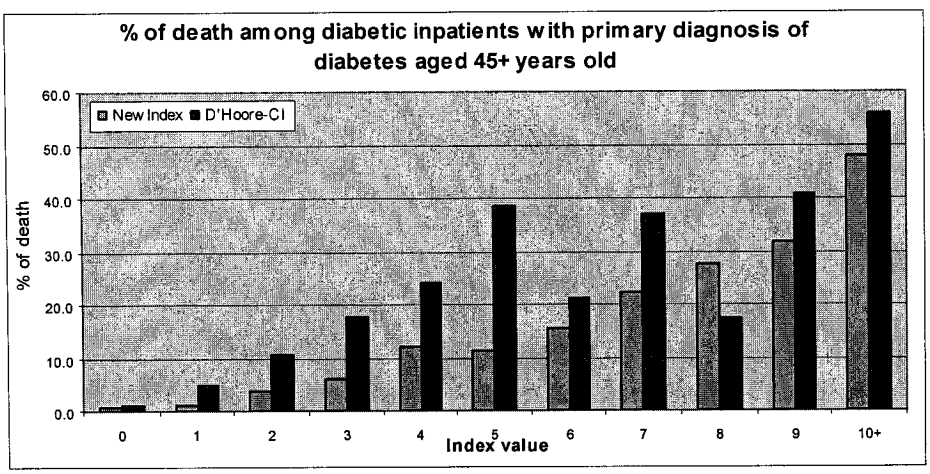


Figure 11

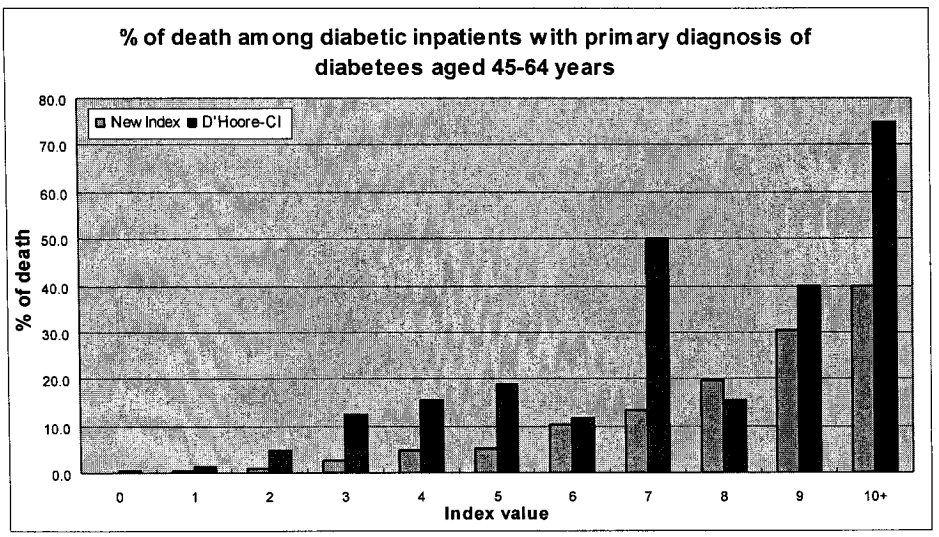
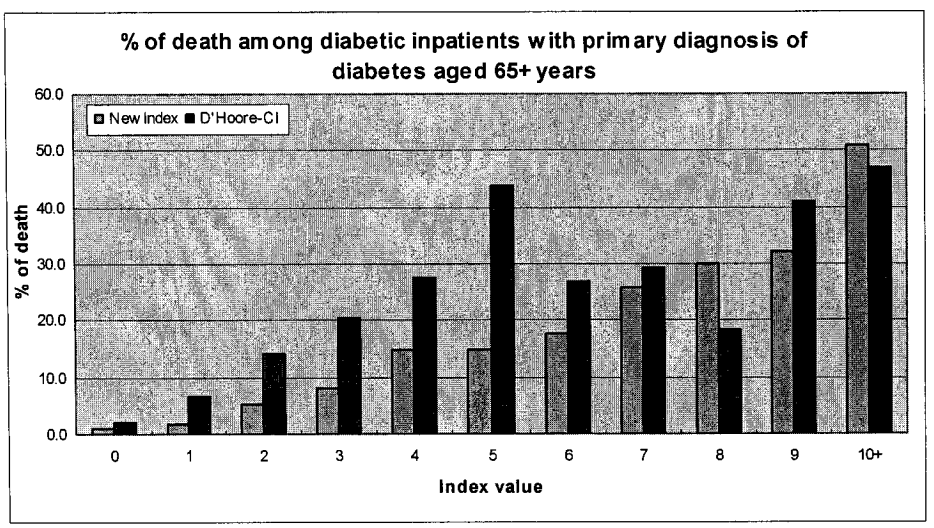


Figure 12



* The testing population consists of a randomly selected episode of hospitalization for each patient in the randomly selected half of the population

5.4.3. Distribution of long length of stay among diabetic patients with primary diagnosis of diabetes in the "testing" population

Table 11 shows some general statistics of length of stay at hospital for inpatients with primary diagnosis of diabetes in the "testing" population. Overall the 75^{th} percentile of length of stay was 11.0 days, which was close to the average stay of 10.4 days. Length of stay $\geq 75^{th}$ percentile was defined as long length of stay.

Figures 13-15 show the distribution of long length of stay by scores of the study-derived index and the D'Hoore-Charlson index. The diabetic patients with a higher score of the study-derived index had higher percentage of long stay at hospital. This linear pattern did not exist for D'Hoore-Charlson scores in all the age groups.

Table 11. Descriptive analysis of length of stay for patients with primary diagnosis of diabetes in the "testing" population* during the period from 1995/96 to 2000/01

	45-64 years (n = 9,903)	≥ 65 years (n = 12,918)	≥ 45 years (n = 22,821)
Mean (days)	8.3	12.0	10.4
25th percentile (days)	3.0	4.0	3.0
Median (days)	5.0	7.0	6.0
75th percentile (days)	9.0	13.0	11.0
% of patients with long LOS	26.8	27.1	27.1

* The testing population consists of a randomly selected episode of hospitalization for each patient in the randomly selected half of the population.

** Longer than or equal to the 75th percentile of length of stay was defined as long length of stay.

Figure 13-15: Crude percentage of long length of stay among diabetic patients with primary diagnosis of diabetes by index values in the "testing" population during the period from 1995/96 to 2000/01

Figure 13

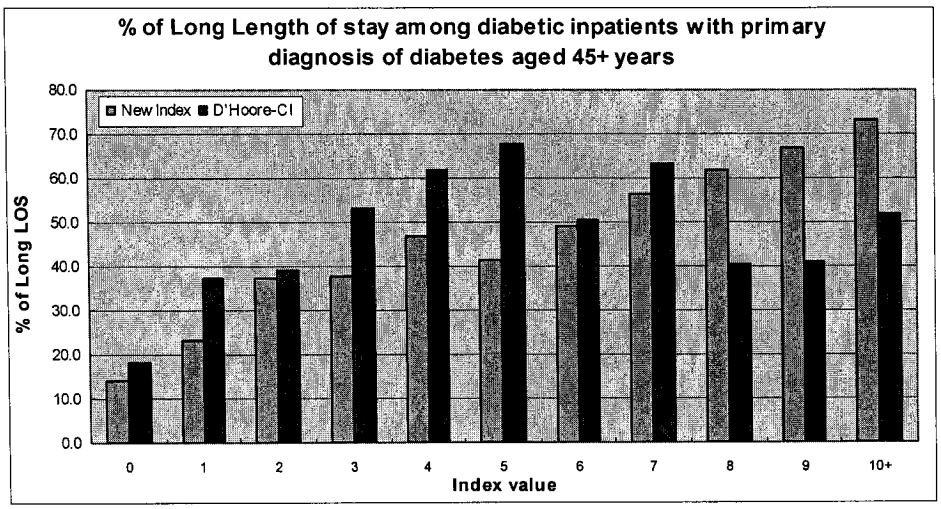


Figure 14

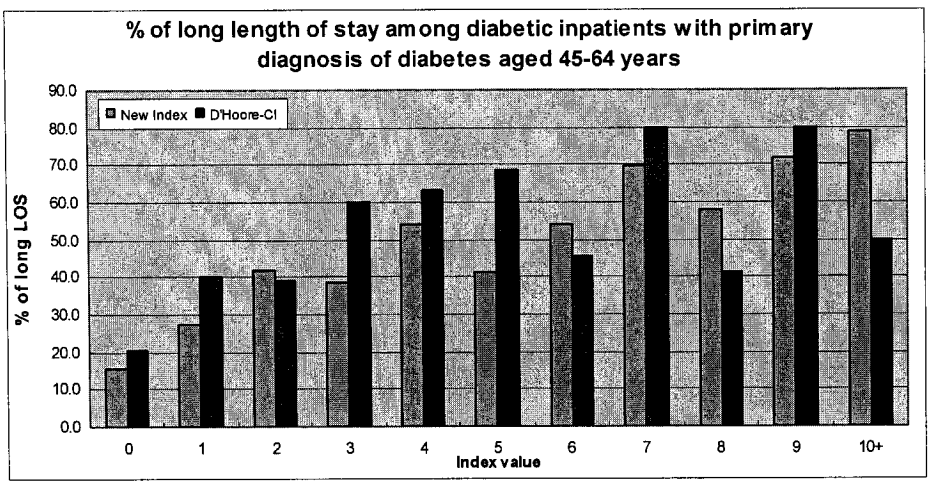
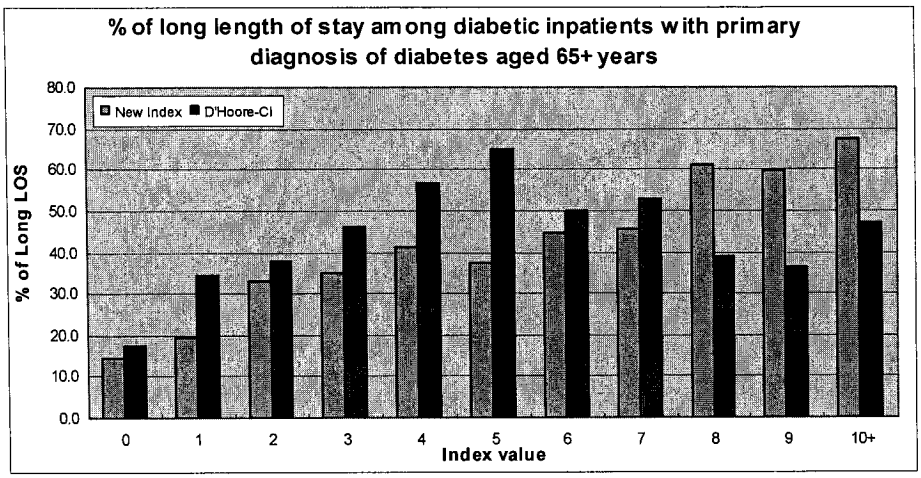


Figure 15



* The Testing population consists of a randomly selected episode of hospitalization for each patient in the randomly selected half of the population

5.4.4. Associations between the indices and outcomes among diabetic patients with primary diagnosis of diabetes in the "testing" population

Table 12 shows the odds ratios for hospital mortality and long length of stay associated with the scores of the study-derived index and D'Hoore-Charlson index among diabetic inpatients with primary diagnosis of diabetes in the "testing" population. In all age groups, the higher the scores of the study-derived index were, the higher the odds ratios for death in hospital were. This trend did not exist for the D'Hoore-Charlson Index.

The scores of the study-derived index and the D'Hoore-Charlson Index were all significantly associated with the odds of staying longer at hospitals in all age groups. Among diabetic patients aged 65 years or over, the higher the scores of the study-derived index were, the higher the odds ratios for staying longer at hospitals were. There was a lack of the consistency for the D'Hoore-Charlson index in all age groups.

Table 12. Adjusted odds ratios and the 95% confidence intervals of hospital mortality and long length of stay for the indices among patients with primary diagnosis of diabetes in the "testing" population* during the period from 1995/96 to 2000/01

	In-hospital Mortality		Long Length of Stay	
	The study-derived Index**	D'Hoore-Charlson Index	The study-derived Index*	D'Hoore-Charlson Index
45-64 years				
0	1.00	1.00	1.00	1.00
1	2.19 (0.84, 5.66)	0.99 (0.98, 1.00)	2.14 (1.87, 2.45)	2.63 (2.30, 3.00)
2	5.10 (2.27, 11.46)	2.03 (2.01, 2.05)	4.01 (3.49, 4.60)	2.48 (2.15, 2.87)
3	11.92 (5.10, 27.82)	11.32 (11.24, 11.40)	3.48 (2.84, 4.26)	5.70 (4.57, 7.12)
4	21.69 (9.79, 48.02)	20.40 (20.20, 20.60)	6.74 (5.34, 8.49)	6.55 (4.56, 9.43)
5	25.69 (11.86, 55.67)	42.71 (41.97, 43.45)	3.76 (2.98, 4.74)	8.52 (4.00, 18.13)
6	61.14 (28.52, 131.04)	8.81 (8.59, 9.03)	6.52 (4.86, 8.75)	3.17 (1.62, 6.22)
7	73.90 (32.91, 165.95)	>999.99	12.17 (8.17, 18.13)	17.05 (3.58, 81.17)
8	116.71 (48.94, 278.34)	20.18 (19.84, 20.52)	7.18 (4.35, 11.83)	2.70 (1.49, 4.90)
9	207.85 (86.07, 501.98)	>999.99	13.42 (7.00, 25.72)	14.40 (1.61, 129.18)
10+	309.14 (147.65, 647.26)	>999.99	20.63 (12.46, 34.14)	3.61 (0.89, 14.70)
≥ 65 years				
0	1.00	1.00	1.00	1.00
1	1.35 (0.88, 2.06)	3.39 (2.69, 4.27)	1.45 (1.27, 1.67)	2.47 (2.23, 2.74)
2	4.48 (3.26, 6.17)	8.47 (6.80, 10.55)	2.97 (2.64, 3.35)	2.93 (2.61, 3.29)
3	6.89 (4.85, 9.81)	13.57 (10.59, 17.39)	3.27 (2.80, 3.83)	4.14 (3.55, 4.84)

	In-hospital Mortality		Long Length of Stay	
	The study-derived Index**	**D'Hoore-Charlson Index**	**The study-derived Index***	**D'Hoore-Charlson Index**
4	14.78 (10.62, 20.55)	21.53 (16.00, 28.96)	4.29 (3.64, 5.05)	6.32 (5.07, 7.87)
5	14.10 (10.04, 19.81)	52.04 (34.11, 79.38)	3.60 (3.02, 4.29)	9.17 (6.20, 13.56)
6	18.73 (13.02, 26.95)	22.20 (12.04, 40.93)	5.04 (4.10, 6.19)	4.98 (2.99, 8.31)
7	29.04 (20.04, 42.09)	22.56 (7.59, 67.06)	5.19 (4.04, 6.49)	5.33 (2.04, 13.91)
8	37.36 (24.85, 56.16)	11.99 (6.88, 20.90)	9.73 (7.25, 13.06)	3.05 (2.02, 4.61)
9	48.27 (30.87, 75.49)	44.20 (18.08, 108.11)	9.24 (6.56, 13.01)	2.96 (1.24, 7.09)
10+	106.08 (72.93, 154.31)	55.84 (20.44, 152.50)	13.06 (9.84, 17.33)	4.37 (1.68, 11.38)
≥ 45 years				
0	1.00	1.00	1.00	1.00
1	1.50 (1.01, 2.20)	1.60 (1.59, 1.61)	1.81 (1.64, 1.99)	2.48 (2.28, 2.69)
2	4.77 (3.55, 6.41)	3.79 (3.78, 3.81)	3.51 (3.20, 3.84)	2.75 (2.51, 3.01)
3	7.76 (5.61, 10.75)	9.90 (9.87, 9.94)	3.50 (3.09, 3.96)	4.78 (4.21, 5.42)
4	16.22 (11.97, 21.98)	21.89 (21.79, 21.99)	5.16 (4.52, 5.89)	6.64 (5.49, 8.03)
5	15.96 (11.69, 21.78)	37.30 (37.00, 37.60)	4.06 (3.53, 4.66)	8.81 (6.19, 12.54)
6	23.65 (17.03, 32.84)	14.75 (14.59, 14.92)	5.75 (4.86, 6.79)	4.49 (2.98, 6.75)
7	34.76 (24.81, 48.72)	40.75 (40.00, 41.51)	7.33 (6.00, 8.96)	7.58 (3.43, 16.74)
8	46.58 (32.17, 67.44)	9.19 (9.10, 9.29)	9.28 (7.19, 11.97)	2.85 (2.03, 4.00)
9	65.29 (43.81, 97.30)	25.72 (25.23, 26.22)	11.46 (8.42, 15.60)	2.85 (1.32, 6.19)
10+	134.91 (96.85, 187.93)	<0.001	16.19 (12.61, 20.78)	4.56 (2.06, 10.07)

* The Testing population consists of a randomly selected episode of hospitalization for each patient in the randomly selected half of the population.

** The study-derived index refers to the index derived from diabetic inpatients aged 45 years or over in the present study.

5.4.5. Predictive performance

Table 12 shows the c statistic and the likelihood ratio chi-square statistic (G^2) of the models for in-hospital mortality and long length of stay among diabetic inpatients with primary diagnosis of diabetes in the "testing" population. In all age groups, the c statistics of the study-derived index in the models for predicting in-hospital mortality were higher than those of the D'Hoore-Charlson Index. Specifically in the age-group of 45-64 years, 65 years or over and 45 years or over they were 0.91 vs. 0.86, 0.85 vs. 0.82, and 0.88 vs. 0.81, respectively. As compared with the baseline regression model, adding either the study-derived index or the D'Hoore-Charlson index improved the fit of the model in all age groups (In the age group of 45+ years, the study-derived index: $\Delta G^2 = 8048.5\text{-}6268.8 = 1779.7$, 10 d.f, $p < 0.001$; the D'Hoore-Charlson index: $\Delta G^2 = 8048.5\text{-}7726.9 = 321.6$, 10 d.f, $p < 0.001$). However, the study-derived index produced G^2 statistics significantly better than the D'Hoore-Charlson index.

The c statistic of the study-derived index in the model for predicting long length of stay among the patients aged 45 years or over was 0.72, which indicates acceptable model discrimination. The c statistics of the study-derived index in the models for predicting long length of stay were higher than those of the D'Hoore-Charlson index in all age groups. The addition of either the study-derived index or the D'Hoore-Charlson index significantly improved the fit of the model in all age groups, compared to the baseline models.

Table 13: Comparison of the study-derived index and the D'Hoore-Charlson index among patients with primary diagnosis of diabetes in the "testing" population* during the period from 1995/96 to 2000/01

Models	45-64 years		≥65 years		≥ 45 years	
	c statistic	G^2 (-2 Log L / d.f.)	c statistic	G^2 (-2 Log L / d.f.)	c statistic	G^2 (-2 Log L / d.f.)
In-hospital Mortality						
Baseline model (Age, sex, province)	0.67	1692.1 / 11	0.65	6339.6 / 11	0.72	8048.5 / 11
Baseline model + the study-derived index**	0.91	1189.7 / 21	0.85	5048.7 / 21	0.88	6268.8 / 21
Baseline model + D'Hoore-CI	0.86	1532.6 / 21	0.82	5351.3 / 21	0.81	7726.9 / 21
Long length of stay						
Baseline model (Age, sex, province)	0.58	11343.7 / 11	0.55	14996.2 / 11	0.61	26012.2 / 11
Baseline model + the study-derived index	0.71	10347.5 / 21	0.69	13865.6 / 21	0.72	23862.2 / 21
Baseline model + D'Hoore-CI	0.67	10751.1 / 21	0.67	14112.9 / 21	0.69	24524.2 / 21

* The Testing population consists of a randomly selected episode of hospitalization for each patient in the randomly selected half of the population.

** The study-derived index refers to the index derived from diabetic inpatients aged 45 years or over in the present study.

6. DISCUSSION

6.1. Summary of Findings

The present study investigated the associations between 48 coexistent general medical conditions and 14 diabetic complications and in-hospital mortality, and found 14 and 8 of them were positively associated death in hospital among diabetic inpatients aged 45 years or over. Of them, respiratory failure, peritonitis and intestinal abscess, and cancer/malignancy including leukemia and lymphoma had a strong association, followed by moderate or severe liver disease, septicemia, acquired immune deficiency syndrome and renal failure. The odds ratios for other important comorbidities ranged from 1.26 to 2.45. The study-derived index included 14 significant coexistent general medical conditions and 8 significant diabetic complications and assigned their odds ratio as their weights for the index.

Some comorbidities had significantly different association with death in hospital in the younger and older age group, and eight comorbidities had stronger association in the younger age group. In spite of these differences, the index developed among diabetic inpatients aged 45 years or over had similar predictive performance as the indices developed in the age-specific groups. Therefore, the present study employed the index developed among diabetic inpatients aged 45 year or over as the study-derived comorbidity index for convenience.

When multivariate analyses were conducted in the 'testing' population, the categorical study-derived comorbidity index, D'Hoore-Charlson Index and number of comorbidities were all significant predictors of in-hospital mortality. The higher the scores of the study-derived index were, the higher the odds ratios for death in hospital

were. This trend was more consistent for the study-derived index compared to either D'Hoore-Charlson index or the number of comorbidities in all age and sex groups.

Some previous studies investigated the impact of multiple comorbidities on health care outcomes among diabetic inpatients. With data drawn from 438 acute care hospitals in California in 1992 (n = 1,779,167), Elixhauser et al. found 13 comorbidities had a significant impact on death in hospital, and the death rate was 7 times higher for diabetic patients with three or more comorbidities compared with those who had no comorbidities (13). However, instead of a single summary index they grouped comorbid conditions into multiple categorical variables. This approach might control for confounding slightly better, but a regression model having large number of independent variables limits its ability to model interactions, decreases its ease of use and decreases its precision in epidemiological analyses (22).

The predictive performance of administrative-based comorbidity indices depends on several factors, including: 1) comorbidities included in an index and their assigned weights; 2) the distribution of comorbid conditions in the source population; 3) the endpoint of a study; and 4) the accuracy of the administrative data. The predictive performance of indices can validly be compared when factors 2-4 are held constant (71). Among diabetic patients aged 45 years or over, the c statistic of the study-derived index was 0.83, which suggests excellent model discrimination (22). The study-derived comorbidity index had better predictive performance than the D'Hoore-Charlson index (c = 0.81) and the number of comorbidities (c = 0.74). The deviance for the D'Hoore-Charlson index or number of comorbidities was much higher as compared with the study-derived index, indicating a worse predictive ability for the D'Hoore-Charlson index or the

number of comorbidities. The differences in the predictive abilities between the study-derived index and the D'Hoore-Charlson index were larger among diabetic patients with primary diagnosis of diabetes. Some investigators have found that comorbidity indices perform best in populations, which are similar to those in which the index was developed(20; 21; 23; 24). They suggested that, in order to increase the predictive power, researchers include in their studies other comorbidities, which are not included in the Charlson index that may be importantly related to the endpoints, and develop and use data-specific empirical weights whenever sufficient representative data are available (20; 21; 23; 24).

The study-derived index improves the predictability over D'Hoore-Charlson Index, but not substantially, among diabetic patients admitted due to any primary diagnosis. The difference is more pronounced among diabetic patients with primary diagnosis of diabetes. The present study could only make efforts on the first factor, "comorbidities included in an index and their assigned weight", to improve the predictive performance of the index. However, to improve predictive performance of such measures requires a combination of strategies: 1) improving the quality of administrative data to have more complete and accurate diagnoses; 2) adding health care utilization and outcome information in administrative data; and 3) linking administrative data with other sources of clinical information. Each of these strategies requires some additional effort and cost (43). With more clinical information such as the severity of primary diagnosis and more accurate and complete coding of comorbidity, the study-derived index may have better predictive performance.

The simplest and most obvious tactic for estimating the overall burden of comorbid diseases would be to find the total number of comorbid diseases for each patient (16). Two studies found that the simple counts of comorbidities performed almost as well as the more complex Charlson scores in prediction of mortality (42; 43). The authors have indicated that the number of diagnoses has the following advantage over an index with assigned weights for comorbidities that: 1) it is easy to calculate and to interpret; 2) it is a complex task to develop a general comorbidity index that can measure both medical resource use and health care outcomes; 3) even if an index is developed that has excellent explanatory and/or predictive ability in one dataset, it is not clear that this will be true for the myriad of datasets that researchers will examine; and 4) Charlson scores may be more susceptible to errors compared to simple counts of the number of diagnoses (42; 43). However, other researchers believe that using the number of comorbidities to predict endpoints does not take into account the seriousness of a comorbid disease, and thus fails to capture important prognostic differences (16). The present study compared the predictive performance of the number of all comorbidities a patient experienced and the study-derived comorbidity index. Although the number of diabetic comorbidities significantly contributed to explain the differences of hospital mortality among diabetic inpatients, the study-derived comorbidity index had much better predictive ability than the number of comorbid diseases, measured by c statistic (0.83 vs. 0.74).

6.2. Strengths of this Study

6.2.1. It was developed and tested on a large, national and more representative population

Based on a large national database which covers almost the entire Canadian population, the present study was conducted in the basis of data from a large national population of all acute hospitals in all ten Canadian provinces that represented diabetic inpatients aged 45 years or over in the general population. The database used in this study includes a unique identifier for each patient, which allowed us to link all hospital stays for each patient and created episodes of hospitalizations by grouping congruent hospital stays together. We then randomly selected one from all episodes of hospitalization during the period from April 1, 1995 to March 31, 2000 for each patient. The resulting index can be applied to health research using large scale administrative data and conducted among diabetic inpatients aged 45 years or over.

6.2.2. More complete and accurate comorbidity profile and outcome reporting compared to simple hospital separation databases

The present study explored all hospital stays of the randomly selected episode of hospitalization for each patient to identify his or her comorbidity profile and the outcome. Thus, the present study is more likely to obtain complete and accurate information about comorbid conditions (independent variables) and discharge status (the outcome) for each diabetic patient, compared to the studies using separation-based data sets. This is particularly true for seriously ill patients who were transferred from one hospital to another.

Furthermore, the data used in the present study provided primary diagnosis and up to 15 other diagnosis codes for each hospital stay except for data from the provinces of Prince Edward Island and Saskatchewan. Many studies have demonstrated that chronic conditions are underreported when discharge abstract forms only allow for the coding of a small number of diagnoses (five diagnoses in Medicare data) (21; 31; 34). Comorbidity is much better reported when the diagnosis field is extended to 9 or 24 diagnoses (66). The value of a comorbidity index based on administrative data largely depends on the completeness and accuracy of diagnostic coding (34). The data used in the present study may provide more complete comorbidity profiles for diabetic patients than the data with only a small number of coding fields, thus, making the study-derived index more valid. When we excluded the two provinces and repeat the analyses, the weights of the comorbidities contained in the indices and the predictive performance of the indices did not significantly change.

6.2.3. The predictive equation was tested on a population that was similar to but distinct from the sample used for development of the equation

The large database allowed us to randomly split the study population in half, one half was used to develop the index, and the other half was used to validate the index. It also allowed us to assess the impact of a broad range of comorbidities on death in hospital while maintaining sufficient statistical power.

6.3. Limitations of the Study

The limitations of the present study are largely attributable to the limitations of administrative data.

6.3.1. Recording bias

In the present study, we found that some comorbidities, such as, hypertension, hypothyroidism/thyroid disease, chronic obstructive pulmonary disease, depression, gastritis/duodenitis, deficiency anemia and gout, had inverse associations with in-hospital mortality, with odds ratios significantly less than 1. This finding is consistent with those from other studies (13; 31; 50) and is an indicator of recording bias in discharge abstract coding that is the severity of the patient's condition inversely affects the coding of chronic unthreatening conditions (13; 63). The possible reasons for this bias have been discussed in the part of the literature review. Physicians are less likely to record chronic conditions in patients who are severely ill or die (13; 25; 27; 50).

6.3.2. Underestimation of comorbid conditions

It is well documented that administrative data tend to underestimate comorbid conditions (22; 23; 31; 71), which may lower the explanatory power of the study-derived index (31). One study compared four populations in different regions of the US with either administrative data or medical records, and striking differences were found in the prevalence of comorbid conditions identified by medical records and administrative data (54). We investigated all hospital stays in the randomly selected episode of hospitalization to identify comorbid condition for each diabetic patient instead of only

using one hospital stay. This method can construct a more complete comorbidity profile for a patient, and may partly overcome the shortcoming of administrative data.

6.3.3. Short-term hospital mortality only

One study found that inpatient death rates depended on length-of-stay patterns and gave a biased picture of mortality (50). Some studies of short-term mortality examined death in 30 days following hospital admission, using the Medicare data files to track patient mortality over time (50; 69), and the researchers believed that observing death in a certain period avoided difficulties in comparing mortality across hospitals caused by differences in hospital discharge practices (50). Further, another study evaluated the relative importance and the explanatory power of D'Hoore-Charlson Index for predicting mortality 30, 90, and 180 days after admission. The study found that the agreement between the model and the data improved as the length of follow-up increased, and the value of the likelihood ratio G^2 statistic and the ability to predict mortality measured by c statistic were greater when modeling 180-day mortality than when modeling 30-day mortality (23).

Our data did not permit identification of out-of-hospital deaths, so we cannot exclude the possibility that the length-of-stay patterns may bias the analyses of in-hospital mortality. However, given that the goal of the present study was to examine differences in likelihood of death related to the presence of given comorbidities at the individual case level (i.e., rather than at hospital level), the issue of differences in hospital discharge practices become less pressing (69). Also, the study-derived index may have a greater explanatory power for predicting long-term mortality among diabetic inpatients, but we cannot confirm this possibility due to lack of the information about out-of-hospital

deaths. It would be expensive and often impossible to follow up individual patients for the information of health care outcomes, especially for large scale of health research. A way to obtain posthospital survival data at a reasonable cost would be to merge large scale discharge abstract data and mortality data.

6.4. The Implication of Comorbidity Index

6.4.1. Health research

Controlling for variation in baseline clinical status including comorbidity is of interest in all types of health care studies. Health status, as measured by disease history, has long been recognized as a major class of potential confounder in most epidemiological studies (22). Policy makers, health authorities and health-service researchers need to compare the effectiveness of different treatments, evaluate health care outcomes achieved by different providers, and investigate the impact of health care policies across different populations (9; 13; 23; 27; 35-38; 76). Health service researchers usually lack the luxury of randomly selecting and allocating patients to treatment protocols, and therefore are often comparing groups having disparate and complex pre-treatment characteristics (23). In addition, eliminating patients with comorbid conditions from studies often results in substantial losses of patients prior to randomization, and limiting the generalizability of the results (77). In each of these health service studies, meaningful comparisons require adjustment for differences in patients' characteristics, which includes the severity of primary diagnosis, sociodemographic factors, functional status and the burden of comorbidity. Since comorbidity may create significant prognostic differences in patients with the same manifestations of the index disease, the impact of

comorbidities on prognosis must be carefully considered when patients are stratified for therapeutic comparisons (11; 38). A comorbidity score, which reflects the aggregate effect of all clinical conditions a patient might have, excluding the disease of primary interest (29), can be a useful tool for controlling for confounding in health research (71).

Some researchers argue that although a comorbidity index is useful because it is easy to use and it saves resources (a major issue when analyzing massive health care databases), adjusting for a score should not be regarded as successfully controlling for confounding for the following reasons. First, it provides only a limited ability to control for confounding (22). Second, a summary score imposes on the analysis a fixed model of the relation between comorbidities and outcome, which is likely to differ among populations (71; 78). Third, a single comorbidity index summarizes a complex construct in an over-simplistic way by making assumptions, which may not be correct (22). Finally, when the outcomes of a particular disease are studied, effects may be underestimated if the disease is a major ingredient of the score. However, the first stage of the studies using large scale administrative databases studies is to digest a vast array of administrative variables into an intelligible and manageable set of proxy variables (22). In this process, the benefits of simplification usually seem much greater than the risks of oversimplification (22). Therefore, a comorbidity index could be a useful tool for preliminary analysis to indicate the direction and magnitude of confounding, which can guide decisions about further analyses (22; 71).

6.4.2. Health care utilization

As the aged population increases in our society, the assessment of comorbidity among patients with diabetes has become progressively important, particularly for health

policy maker and administrators in regard to health planning. The present study found among diabetic inpatients aged 45 years or over, only 11.4% did not have any comorbidity, and the odds of death in hospital among those with the study-derived score of 3 was over 7 times higher than that among those with the score of zero. Therefore, diabetic comorbidities including diabetic complications and coexistent medical conditions are critical issues in diabetes prevention and control. Simply ignoring comorbidity may create ineffectiveness in delivery of care that is greater than any gains made through management of a single disease (79). Health plan administrators may benefit from a validated comorbidity index that could help evaluate the health care outcomes among diabetic patients when other conditions co-exist. Health care outcomes may be improved if diabetes disease management that address the reality of multiple related and unrelated comorbidities can successfully be aimed at preventing diabetic complications and controlling comorbidities (79).

Furthermore, The National Diabetes Surveillance System (NDSS) is planning to investigate the burden of comorbidity in peoples with diabetes. A comorbidity index derived from diabetic population may provide this survey useful information about which kind of comorbidities have significant impact on health care outcomes among people with diabetes.

6.4.3. Clinical treatment

A diabetes comorbidity index may be useful particularly to clinicians caring for diabetic patients with complex medical profiles, including the aged and those with other chronic diseases (19). Compared to patients who do not have comorbidity, patients with comorbid illness may be at greater risk of death, less able to tolerate particular medical

procedures, and less responsive to therapy (19). For example, studies involving breast and prostate cancer patients have found that patients with more comorbidities receive less aggressive treatment for their tumors, even after controlling for patients' age and cancer stage (80; 81). The choice and goals of therapy might be affected by such patient's characteristics as advanced age, the presence of comorbid conditions, and the patient's capacity to understand (82). With the understanding of the impact of comorbidities on death in hospital and the score particularly developed for diabetic inpatients, physicians may roughly quantify the coexistent diseases, estimate the prognostic of a patient and factor the presence of particular comorbidities into decisions concerning the most appropriate medical treatments for patients (27; 83).

6.5. Future studies

A major criticism of all comorbidity adjustment methods is that they were developed to predict a particular type of outcome (e.g. morbidity or mortality) but are used to adjust for the risk of other outcomes (e.g. health service use or costs). The original Charlson index was developed for predicting one-year all cause mortality among patients admitted in New York Hospital-Cornell Medical Center during 1-month period in 1984 and validated among women with breast cancer. It has been widely used to predict various kinds of outcomes including health care utilization such as length of stay, readmission and hospital charges, and health care outcomes such as post-operation complication and death among different populations with medical conditions or undergoing surgeries (20; 21; 24; 31-35). Although there are common cases when the severity of one outcome is inversely related to the intensity of health care use (e.g. sudden cardiac death), in the aggregate, adverse outcomes are positively correlated with each

other (22). Therefore, a comorbidity score developed in one setting can be applied in a very different setting as long as it is only for exploratory purposes (22).

The study-derived index was developed from Canadian diabetic inpatients aged 45 years or over admitted due to any primary diagnosis. This study is the first of which we are aware to develop an index especially for diabetic inpatients to examine the impact of not only diabetic complications but co-existent general medical conditions. The c statistic of 0.83 achieved with the study-derived index suggests that the index provide substantial independent information for predicting in-hospital mortality even after adjusting for demographic factors. The predictive performance of the study-derived index was superior to that of the D'Hoore-Charlson index. Among diabetic patients admitted to hospitals for diabetes, the models incorporating the study-derived index produced acceptable model discrimination for predicting the odds of staying longer at hospital. We invite researchers to further validate the study-derived index among diabetic inpatients with broader range of age or for predicting other outcomes such as readmission and length of stay.

In a future study, we plan to link the Health Person Oriented Information database to the Mortality database to track posthospital status in order to predict long-term mortality using comorbidity among diabetic population.

6.6. Conclusion

In conclusion, the present study identified 14 general medical conditions and 8 groups of diabetic complications that were positively associated with death in hospital among diabetic inpatients aged 45 years or over. The predictive performance of the study-derived index was superior to those of D'Hoore-Charlson index and the number of

comorbidities. The study-derived index can be a useful tool to control for potential confounding caused by comorbidity in the exploratory data analysis of diabetes research, to assist in creating more effective diabetes disease management systems, and to estimate the prognostic of diabetic inpatients for health care provider.

REFERENCE

1. Health Canada. Diabetes in Canada, 2nd ed. Center for Chronic Disease Prevention and Control, Population and Public Health Branch, Health Canada, 2002:27; 45.

2. Gu K, Cowie CC, Harris MI: Diabetes and decline in heart disease mortality in US adults. *Jama* 281:1291-1297, 1999

3. Health Canada. Diabetes in Canada: national statistics and opportunities for improved surveillance, prevention, and control. Ottawa: Health Canada, 1999:35.

4. Ray NF, Thamer M, Taylor T, Fehrenbach SN, Ratner R: Hospitalization and expenditures for the treatment of general medical conditions among the U.S. diabetic population in 1991. *J Clin Endocrinol Metab* 81:3671-3679, 1996

5. Laing SP, Swerdlow AJ, Slater SD, Botha JL, Burden AC, Waugh NR, Smith AW, Hill RD, Bingley PJ, Patterson CC, Qiao Z, Keen H: The British Diabetic Association Cohort Study, II: cause-specific mortality in patients with insulin-treated diabetes mellitus. *Diabet Med* 16:466-471, 1999

6. National Institute of Diabetes and Digestive and Kidney Disease. Diabetes in America. 2nd ed. chapter 11. Washington DC: National Institute of Health; 1995. NIH publication 95-1468. 1995

7. Saydah SH, Eberhardt MS, Loria CM, Brancati FL: Age and the burden of death attributable to diabetes in the United States. *Am J Epidemiol* 156:714-719, 2002

8. Health Canada. Diabetes in Canada, 2nd ed. Center for Chronic Disease Prevention and Control, Population and Public Health Branch, Health Canada, 2002:58.

9. van Doorn C, Bogardus ST, Williams CS, Concato J, Towle VR, Inouye SK: Risk adjustment for older hospitalized persons: a comparison of two methods of data collection for the Charlson index. *Journal of Clinical Epidemiology* 54:694-701, 2001

10. Kozma CM: An opportunity for increased efficiency. *Manag Care Interface* 15:46-47, 2002

11. Feinstein AR: The pre-therapeutic classification of comorbidity in chronic disease. *Journal of Chronic Diseases* 23:455-469, 1970

12. American Diabetes Association. Economic consequences of diabetes mellitus in the U.S. in 1997. *Diabetes Care* 21:296-309, 1998

13. Elixhauser A, Steiner C, Harris DR, Coffey RM: Comorbidity measures for use with administrative data. *Med Care* 36:8-27, 1998

14. Pascual JM, Gonzalez C, de Juan S, Sanchez C, Sanchez B, Perez M: Impact of diabetes mellitus on hospitalization costs. *Med Clin (Barc)* 107:207-210, 1996

15. Greenfield S, Aronow HU, Elashoff RM, Watanabe D: Flaws in mortality data. The hazards of ignoring comorbid disease. *Jama* 260:2253-2255, 1988

16. Charlson ME, Pompei P, Ales KL, MacKenzie CR: A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 40:373-383, 1987

17. Clark DO, Von Korff M, Saunders K, Baluch WM, Simon GE: A chronic disease score with empirically derived weights. *Med Care* 33:783-795, 1995

18. Von Korff M, Wagner EH, Saunders K: A chronic disease score from automated pharmacy data. *J Clin Epidemiol* 45:197-203, 1992

19. Klabunde CN, Potosky AL, Legler JM, Warren JL: Development of a comorbidity index using physician claims data. *Journal of Clinical Epidemiology* 53:1258-1267, 2000

20. Ghali WA, Hall RE, Rosen AK, Ash AS, Moskowitz MA: Searching for an improved clinical comorbidity index for use with ICD-9-CM administrative data. *Journal of Clinical Epidemiology* 49:273-278, 1996

21. Romano PS, Roos LL, Jollis JG: Further evidence concerning the use of a clinical comorbidity index with ICD-9-CM administrative data. *J Clin Epidemiol* 46:1085-1090, 1993

22. Schneeweiss S, Maclure M: Use of comorbidity scores for control of confounding in studies using administrative databases. *Int J Epidemiol* 29:891-898, 2000

23. Cleves MA, Sanchez N, Draheim M: Evaluation of two competing methods for calculating Charlson's comorbidity index when analyzing short-term mortality using administrative data. *Journal of Clinical Epidemiology* 50:903-908, 1997

24. Romano PS, Roos LL, Jollis JG: Adapting a clinical comorbidity index for use with ICD-9-CM administrative data: differing perspectives. *Journal of Clinical Epidemiology* 46:1075-1079; discussion 1081-1090, 1993

25. Iezzoni LI, Foley SM, Daley J, Hughes J, Fisher ES, Heeren T: Comorbidities, complications, and coding bias. Does the number of diagnosis codes matter in predicting in-hospital mortality? *Jama* 267:2197-2203, 1992

26. Shwartz M, Iezzoni LI, Moskowitz MA, Ash AS, Sawitz E: The importance of comorbidities in explaining differences in patient costs. *Med Care* 34:767-782, 1996

27. Rochon PA, Katz JN, Morrow LA, McGlinchey-Berroth R, Ahlquist MM, Sarkarati M, Minaker KL: Comorbid illness is associated with survival and length of hospital stay

in patients with chronic disability. A prospective comparison of three comorbidity indices. *Medical Care* 34:1093-1101, 1996

28. Powell H, Lim LL, Heller RF: Accuracy of administrative data to assess comorbidity in patients with heart disease. an Australian perspective. *J Clin Epidemiol* 54:687-693, 2001

29. Greenfield S, Nelson EC: Recent developments and future issues in the use of health status assessment measures in clinical settings. *Med Care* 30:MS23-41, 1992

30. Druss BG, Marcus SC, Olfson M, Tanielian T, Elinson L, Pincus HA: Comparing the national economic burden of five chronic conditions. *Health Aff (Millwood)* 20:233-241, 2001

31. D'Hoore W, Bouckaert A, Tilquin C: Practical considerations on the use of the Charlson comorbidity index with administrative data bases. *Journal of Clinical Epidemiology* 49:1429-1433, 1996

32. Roos LL, Sharp SM, Cohen MM, Wajda A: Risk adjustment in claims-based research: the search for efficient approaches. *J Clin Epidemiol* 42:1193-1206, 1989

33. Roos LL, Stranc L, James RC, Li J: Complications, comorbidities, and mortality: improving classification and prediction. *Health Serv Res* 32:229-238; discussion 239-242, 1997

34. Deyo RA, Cherkin DC, Ciol MA: Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *Journal of Clinical Epidemiology* 45:613-619, 1992

35. D'Hoore W, Sicotte C, Tilquin C: Risk adjustment in outcome assessment: the Charlson comorbidity index. *Methods of Information in Medicine* 32:382-387, 1993

36. Stukenborg GJ, Wagner DP, Connors AF, Jr.: Comparison of the performance of two comorbidity measures, with and without information from prior hospitalizations. *Med Care* 39:727-739, 2001

37. Kieszak SM, Flanders WD, Kosinski AS, Shipp CC, Karp H: A comparison of the Charlson comorbidity index derived from medical record data and administrative billing data. *Journal of Clinical Epidemiology* 52:137-142, 1999

38. Librero J, Peiro S, Ordinana R: Chronic comorbidity and outcomes of hospital care: length of stay, mortality, and readmission at 30 and 365 days. *J Clin Epidemiol* 52:171-179, 1999

39. El-Kebbi IM, Ziemer DC, Cook CB, Miller CD, Gallina DL, Phillips LS: Comorbidity and glycemic control in patients with type 2 diabetes. *Arch Intern Med* 161:1295-1300, 2001

40. Glynn RJ, Monane M, Gurwitz JH, Choodnovskiy I, Avorn J: Agreement between drug treatment data and a discharge diagnosis of diabetes mellitus in the elderly. *Am J Epidemiol* 149:541-549, 1999

41. Redelmeier DA, Tan SH, Booth GL: The treatment of unrelated disorders in patients with chronic medical diseases. *N Engl J Med* 338:1516-1520, 1998

42. Melfi C, Holleman E, Arthur D, Katz B: Selecting a patient characteristics index for the prediction of medical outcomes using administrative claims data. *J Clin Epidemiol* 48:917-926, 1995

43. Wang PS, Walker A, Tsuang M, Orav EJ, Levin R, Avorn J: Strategies for improving comorbidity measures based on Medicare and Medicaid claims data. *J Clin Epidemiol* 53:571-578, 2000

44. Satariano WA, Ragheb N, Dupuis M: Comorbidity in older women with breast cancer: An epidemiologic approach. In *Cancer in the Elderly: Approaches to Early Diagnosis and Treatment* New York, Springer Press, 1989, p. 71-107

45. Malenka DJ, McLerran D, Roos N, Fisher ES, Wennberg JE: Using administrative data to describe casemix: a comparison with the medical record. *J Clin Epidemiol* 47:1027-1032, 1994

46. Siegler EL, Stineman MG, Maislin G: Development of complications during rehabilitation. *Arch Intern Med* 154:2185-2190, 1994

47. Kaplan MH, Feinstein AR: The importance of classifying initial co-morbidity in evaluatin the outcome of diabetes mellitus. *J Chronic Dis* 27:387-404, 1974

48. Knaus WA, Wagner DP, Draper EA, Zimmerman JE, Bergner M, Bastos PG, Sirio CA, Murphy DJ, Lotring T, Damiano A, et al.: The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. *Chest* 100:1619-1636, 1991

49. Knaus WA, Draper EA, Wagner DP, Zimmerman JE: APACHE II: a severity of disease classification system. *Crit Care Med* 13:818-829, 1985

50. Jencks SF, Williams DK, Kay TL: Assessing hospital-associated deaths from discharge data. The role of length of stay and comorbidities. *Jama* 260:2240-2246, 1988

51. Hannan EL, Kilburn H, Jr., Lindsey ML, Lewis R: Clinical versus administrative data bases for CABG surgery. Does it matter? *Med Care* 30:892-907, 1992

52. Newton KM, Wagner EH, Ramsey SD, McCulloch D, Evans R, Sandhu N, Davis C: The use of automated data to identify complications and comorbidities of diabetes: a validation study. *J Clin Epidemiol* 52:199-207, 1999

53. Hawker GA, Coyte PC, Wright JG, Paul JE, Bombardier C: Accuracy of administrative data for assessing outcomes after knee replacement surgery. *J Clin Epidemiol* 50:265-273, 1997

54. Romano PS, Roos LL, Luft HS, Jollis JG, Doliszny K: A comparison of administrative versus clinical data: coronary artery bypass surgery as an example. Ischemic Heart Disease Patient Outcomes Research Team. *J Clin Epidemiol* 47:249-260, 1994

55. Charlson ME, Ales KL, Simon R, MacKenzie CR: Why predictive indexes perform less well in validation studies. Is it magic or methods? *Arch Intern Med* 147:2155-2161, 1987

56. Roos NP, Wennberg JE, Malenka DJ, Fisher ES, McPherson K, Andersen TF, Cohen MM, Ramsey E: Mortality and reoperation after open and transurethral resection of the prostate for benign prostatic hyperplasia. *N Engl J Med* 320:1120-1124, 1989

57. Malenka DJ, Roos N, Fisher ES, McLerran D, Whaley FS, Barry MJ, Bruskewitz R, Wennberg JE: Further study of the increased mortality following transurethral prostatectomy: a chart-based analysis. *J Urol* 144:224-227; discussion 228, 1990

58. Concato J, Horwitz RI, Feinstein AR, Elmore JG, Schiff SF: Problems of comorbidity in mortality after prostatectomy. *Jama* 267:1077-1082, 1992

59. Morris JA, Jr., MacKenzie EJ, Edelstein SL: The effect of preexisting conditions on mortality in trauma patients. *Jama* 263:1942-1946, 1990

60. Iezzoni LI, Ash AS, Coffman GA, Moskowitz MA: Predicting in-hospital mortality. A comparison of severity measurement approaches. *Med Care* 30:347-359, 1992

61. Green J, Wintfeld N: How accurate are hospital discharge data for evaluating effectiveness of care? *Med Care* 31:719-731, 1993

62. Normand SL, Morris CN, Fung KS, McNeil BJ, Epstein AM: Development and validation of a claims based index for adjusting for risk of mortality: the case of acute myocardial infarction. *J Clin Epidemiol* 48:229-243, 1995

63. Hughes JS, Iezzoni LI, Daley J, Greenberg L: How severity measures rate hospitalized patients. *J Gen Intern Med* 11:303-311, 1996

64. Steiner CA, Bass EB, Talamini MA, Pitt HA, Steinberg EP: Surgical rates and operative mortality for open and laparoscopic cholecystectomy in Maryland. *N Engl J Med* 330:403-408, 1994

65. Glynn RJ, Monane M, Gurwitz JH, Choodnovskiy I, Avorn J: Aging, comorbidity, and reduced rates of drug treatment for diabetes mellitus. *J Clin Epidemiol* 52:781-790, 1999

66. Romano PS, Mark DH: Bias in the coding of hospital discharge data and its implications for quality assessment. *Med Care* 32:81-90, 1994

67. Iezzoni LI, Daley J, Heeren T, Foley SM, Fisher ES, Duncan C, Hughes JS, Coffman GA: Identifying complications of care using administrative data. *Med Care* 32:700-715, 1994

68. Iezzoni LI: Using administrative diagnostic data to assess the quality of hospital care. Pitfalls and potential of ICD-9-CM. *Int J Technol Assess Health Care* 6:272-281, 1990

69. Iezzoni LI, Heeren T, Foley SM, Daley J, Hughes J, Coffman GA: Chronic conditions and risk of in-hospital death. *Health Serv Res* 29:435-460, 1994

70. Newschaffer CJ, Bush TL, Penberthy LT: Comorbidity measurement in elderly female breast cancer patients with administrative and medical records data. *J Clin Epidemiol* 50:725-733, 1997

71. Schneeweiss S, Seeger JD, Maclure M, Wang PS, Avorn J, Glynn RJ: Performance of comorbidity scores to control for confounding in epidemiologic studies using claims data. *Am J Epidemiol* 154:854-864, 2001

72. Fried L, Bernardini J, Piraino B: Charlson comorbidity index as a predictor of outcomes in incident peritoneal dialysis patients. *American Journal of Kidney Diseases* 37:337-342, 2001

73. Statistic Canada. Hospital Person Oriented Information Databse (HPOI) data dictionary.Health statistics division, Statistic Canada, 2004.

74. American Diabetes Association. Economic cost of diabetes in the U.S. in 2002: Diabetes Care 26:917-932, 2003.

75. Hosmer D, Lemeshow S: *Applied logistic regression.* New York, John Wiley and Sons, 2000

76. Cornoni-Huntley JC, Foley DJ, Guralnik JM: Co-morbidity analysis: a strategy for understanding mortality, disability and use of health care facilities of older people. *Int J Epidemiol* 20 Suppl 1:S8-17, 1991

77. Charlson ME, Horwitz RI: Applying results of randomised trials to clinical practice: impact of losses before randomisation. *Br Med J (Clin Res Ed)* 289:1281-1284, 1984

78. Katz D, Foxman B: How well do prediction equations predict? Using receiver operating characteristic curves and accuracy curves to compare validity and generalizability. *Epidemiology* 4:319-326, 1993

79. Garis RI, Shara MA, Farmer KC, Horrell JF, Arora M: The cost of diabetes in the presence of comorbid conditions. *Manag Care Interface* 15:48-53, 55, 2002

80. Bennett CL, Greenfield S, Aronow H, Ganz P, Vogelzang NJ, Elashoff RM: Patterns of care related to age of men with prostate cancer. *Cancer* 67:2633-2641, 1991

81. Greenfield S, Blanco DM, Elashoff RM, Ganz PA: Patterns of care related to age of breast cancer patients. *Jama* 257:2766-2770, 1987

82. American Diabetes Association, Standards of medical care for patients with diabetes mellitus. Diabetes Care 17 (1994), pp. 616-623.

83. Reid BC, Alberg AJ, Klassen AC, Rozier RG, Garcia I, Winn DM, Samet JM: A comparison of three comorbidity indexes in a head and neck cancer population. O*ral Oncol* 38:187-194, 2002

Appendix 1: D'Hoore-Charlson Index

Weights	Conditions	ICD-9 codes
1	Myocardial infarction	410, 411
	Congestive heart failure	398, 402, 428
	Peripheral vascular disease	440-447
	Dementia	290, 291, 294
	Cerebrovascular disease	430-433, 435
	Chronic pulmonary disease	491-493
	Connective tissue disease	710, 714, 725
	Ulcer disease	531-534
	Mild liver disease	571, 573
2	Hemiplegia	342, 434, 436, 437
	Moderate or severe renal disease	403, 404, 580-586
	Diabetes	250
	Any tumor	140-195
	Leukemia	204-208
	Lymphoma	200, 202, 203
3	Moderate or severe liver disease	070, 570, 572
6	Metastatic solid tumor	196-199

Appendix 2. The initial list of diabetic comorbidities

Comorbidity	ICD-9 codes
Cardiac valve disease	394-396, 397, 424
Hypertension	401-405
Pulmonary circulation disorders	415, 416, 417
Chronic pulmonary disease	491-493
Respiratory failure, insufficiency, arrest (adult)	518.5, 518.8, 799.1, V46.1
Lower respiratory disease (excluding asthma, pleurisy, and respiratory failure)	513-516
Asthma	493
Tuberculosis	010 - 018
Hypothyroidism / Thyroid disease	240 - 246
Mild liver disease	571, 573
Moderate or severe liver disease	070, 570, 572
Ulcer disease	531 - 534
Acquired immune deficiency syndrome (AIDS)	042, 043, 044
Cancer/ malignancy (incl. leukaemia & lymphoma)	140-195, 196-198, 199, 200-208, 230-234, 235-238, 239
Coagulopathy	286, 287.1, 287.3-287.5,
Obesity	278.0
Fluid and electrolyte disorders	276
Blood loss anemia	280.0

Deficiency anemia	280.1, 280.8, 280.9, 281, 285.9
Alcohol abuse	291, 303, 305.0, V11.3
Drug abuse	292, 304, 305
Psychoses	295-298, 299.1
Depression	300.4, 309.1, 311
Affective disorders	296, 298.0, 301.1,
Cystic fibrosis	277.0
Peritonitis and intestinal abscess	567, 569.5
Male genital disorders (excluding hyperplasia of prostate and inflammatory conditions)	602, 603, 605 - 608, 792.2
Septicemia (excluding during labor)	003.1, 020.0, 022.3, 036.2, 038, 054.5, 790.7
Gastritis and duodenitis, Other and unspecified noninfectious gastroenteritis and colitis	535, 558.9
Crohn's disease, colitis	555.9
Fracture of neck of femur (hip)	820, 905.3
Hereditary and degenerative nervous system disorders (excluding Parkinson's and multiple sclerosis)	330, 331.3, 331.4, 331.7-331.9, 333.0-333.8, 334, 335, 336, 337.0, 337.1, 337.9
Parkinson's disease	332
Epilepsy	345.9
Gout	274
Connective tissue disease/ Rheumatoid arthritis/Collagen vascular disease	710, 714, 725, 701.0, 720
Infections of kidney, Cystitis	590, 595
Other disorders of bladder, Renal sclerosis (unspecified)	596, 587

Urinary tract infection	599.0
Hypertriglyceridemia, other and unspecified hyperlipidemia, mixed hyperlipidemia	272.1, 272.4, 272.2
Other specified endocrine disorders	259.8
Other retinal disorders, Cataract, Glaucoma, Visual disturbance, low vision, blindness	362, 365, 366, 368-369
Disorders of the optic nerve and visual pathways	377
Infective otitis externa	380.1
Degenerative skin disorders	709.3
Candidiasis of vulva and vagina	112.1

Appendix 3. Diabetic complications (analysis groups)

Chronic complications of diabetes	ICD-9 codes
Group 1	
Other specified idiopathic peripheral neuropathy	356.8
Mononeuritis of upper and lower limbs	354, 355
Peripheral autonomic neuropathy	337.1
Polyneuropathy in diabetes	357.2
Neuralgia, neuritis, and radiculitis, unspecified	729.2
Diabetes with neurological complications	250.6
Group 2	
Occlusion of cerebral arteries	434
Late effects of cerebrovascular disease	438
Occlusion of stenosis of pre-cerebral arteries	433
Other and ill-defined cerebrovascular disease	437
Acute, but ill-defined, cerebrovascular disease	436
Transient ischemic attack	435
Group 3	
Arthropathy associated w/neurological disorders (Charcot's arthropathy)	713.5
Group 4	
Embolism and thrombosis, structure of artery	444, 447.1
Phlebitis and thrombophlebitis, portal vein thrombosis, and thrombolism and venous thrombolism	451, 452
Other venous embolism and thrombolism	453

Group 5	
Atherosclerosis	440
Other peripheral vascular disease	443
Other disorders of circulatory system	459
Gangrene and amputations	785.4, 885–887, 895–897
Diabetes w/peripheral circulatory disorders	250.7
Group 6	
Chronic ulcer of skin	707
Cellulitis	681, 682
Group 7A	
Myocardial infarction	410, 412
Group 7B	
Angina	413
Atherosclerotic cardiovascular disease	429.2
Other acute and subacute forms of ischemic heart disease	411
Other chronic ischemic heart disease	414
Group 8	
Cardiomegaly	429.3
Cardiomyopathy	425
Heart failure	428
Myocardial degeneration	429.1
Group 9	

Conduction disorders and cardiac dysrhythmias	426–427
Group 10A	
Nephrotic syndrome, nephropathy	580–583
Proteinuria	791.0
Diabetes and renal complications	250.4
Group 10B	
Renal failure and its sequelae	584, 586, 588
Chronic renal failure (end-stage renal disease)	585
Group 11	
Diabetes with ophthalmic complications	250.5
Group 12	
Diabetes with other specified manifestations	250.8
Diabetes with unspecified complication	250.9

Appendix 4. Adjusted odds ratios and the 95% confidence intervals for hospital mortality associated with the study derived index* in the "testing" population during the period from 1995/96 to 2000/01**

	Male	Female	Total
45-64 years			
0	1.00	1.00	1.00
1	1.67 (1.11, 2.56)	2.43 (1.52, 3.88)	1.97 (1.45, 2.69)
2	4.88 (3.91, 6.10)	7.92 (6.02, 10.42)	5.91 (4.97, 7.03)
3	8.39 (6.29, 11.18)	13.49 (9.51, 19.13)	10.17 (8.14, 12.70)
4	13.78 (10.95, 17.35)	22.67 (17.05, 30.15)	16.74 (14.00, 20.03)
5	34.34 (27.94, 42.20)	37.61 (29.33, 48.25)	35.76 (30.51, 41.91)
6	46.03 (36.86, 57.48)	53.48 (40.39, 70.83)	49.40 (41.51, 58.80)
7	50.04 (39.48, 63.41)	80.78 (60.23, 108.33)	60.59 (50.39, 72.86)
8	62.22 (47.94, 80.75)	99.41 (72.50, 136.33)	75.26 (61.56, 92.01)
9	119.68 (92.54, 154.80)	144.35 (103.96, 200.43)	130.19 (106.35, 159.38)
10+	197.96 (159.47, 245.75)	241.20 (184.09, 316.01)	216.25 (182.67, 256.00)
≥ 65 years			
0	1.00	1.00	1.00
1	2.07 (1.79, 2.40)	1.71 (1.48, 1.97)	1.89 (1.68, 2.06)
2	3.91 (3.57, 4.30)	4.85 (4.43, 5.31)	4.38 (4.10, 4.67)
3	6.24 (5.57, 7.00)	6.51 (5.81, 7.28)	6.39 (5.90, 6.92)
4	8.06 (7.32, 8.80)	9.91 (9.02, 10.89)	8.98 (8.39, 9.60)
5	11.54 (10.53, 12.66)	11.85 (10.80, 13.01)	11.77 (11.03, 12.57)

	Male	Female	Total
6	15.77 (14.25, 17.45)	18.77 (16.96, 20.78)	17.23 (16.04, 18.51)
7	17.21 (15.53, 19.06)	19.91 (17.92, 22.12)	18.54 (17.23, 19.95)
8	23.98 (21.38, 26.89)	28.68 (25.46, 32.32)	26.24 (24.16, 28.50)
9	26.60 (23.64, 29.93)	29.58 (26.00, 33.64)	28.15 (25.81, 30.70)
10+	51.44 (46.56, 56.83)	57.13 (51.47, 63.41)	54.45 (50.67, 58.50)
≥ 45 years			
0	1.00	1.00	1.00
1	2.11 (1.84, 2.42)	1.87 (1.63, 2.15)	1.97 (1.79, 2.17)
2	4.20 (3.85, 4.58)	5.36 (4.92, 5.84)	4.74 (4.46, 5.04)
3	6.87 (6.18, 7.64)	7.40 (6.65, 8.23)	7.13 (6.62, 7.69)
4	9.18 (8.40, 10.03)	11.35 (10.37, 12.41)	10.19 (9.57, 10.86)
5	14.47 (13.31, 15.74)	14.21 (13.03, 15.49)	14.42 (13.58, 15.32)
6	19.70 (17.97, 21.60)	21.95 (19.95, 24.16)	20.79 (19.46, 22.21)
7	21.21 (19.30, 23.30)	24.15 (21.87, 26.67)	22.60 (21.11, 24.20)
8	29.06 (26.16, 32.28)	34.38 (30.75, 38.45)	31.51 (29.19, 34.01)
9	35.65 (32.01, 39.71)	37.30 (33.08, 42.06)	36.65 (33.83, 39.70)
10+	68.81 (62.87, 75.32)	71.55 (64.94, 78.84)	70.59 (66.08, 75.40)

* The study-derived index refers to the index derived from diabetic inpatients aged 45 years or over in the present study.

** The Testing population consists of a randomly selected episode of hospitalization for each patient in the randomly selected half of the population