CREATING A NATIONAL REGISTER OF CHILDHOOD TYPE 1 DIABETES USING ROUTINELY COLLECTED HOSPITAL DATA

Susan Hodgson[1,2], Linda Beale[2], Roger C Parslow[3], Richard G Feltbower[3], Lars Jarup[4]

[1] Institute of Health and Society, Newcastle University, Newcastle upon Tyne, NE1 4HH, UK

[2] MRC-HPA Centre for Environment and Health, Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, W2 1PG, UK

[3] Paediatric Epidemiology Group, University of Leeds, Leeds, LS2 9JT, UK

[4] Small Area Health Statistics Unit, Imperial College London, W2 1PG, UK

Word count for entire manuscript: 5,862

Abbreviations:

CI	Confidence Interval
HES	Hospital Episodes Statistics
ICD	International Classification of Disease
NHS	National Health Service
SAHSU	Small Area Health Statistics Unit

ABSTRACT:

Introduction: There is no national register of childhood type 1 diabetes mellitus for England. Our aim was to assess the feasibility of using routine hospital admissions data as a surrogate for a childhood diabetes register across England, and to create a geographically referenced childhood diabetes dataset for use in epidemiologic studies and health service research.

Methods: Hospital Episodes Statistics data for England from April 1992 to March 2006 referring to a type 1 diabetes diagnosis in 0-14 year olds were cleaned to approximate an incident dataset. The cleaned data were validated against regional population based register data, available for Yorkshire and the area of the former Oxford Regional Health Authority.

Results: There were 32,665 unique cases of type 1 and type unknown diabetes over the study period. The hospital derived data improved in quality over time (91% concordance with regional register data over the period 2000-2006, versus 52% concordance over the period 1992-1999), and data quality was better for younger (0-9 years) (86.5% concordance with regional register data) than older cases (10-14 years). Overall incidence was 24.99 (95% Confidence Interval 24.71-25.26) per 100,000. Basic trends in age distribution, seasonality of onset, and incidence matched well with previously reported findings.

Conclusion: We were able to create a surrogate register of childhood diabetes based on national hospital admissions data, containing ~2,300 cases per year, and geo-coded to a high resolution. For younger cases (0-9 years) and more recent years (from 2000) these data will be a useful resource for epidemiological studies exploring the determinants of childhood diabetes.

Keywords:

Diabetes Mellitus, Type 1; Registries; Hospital Records; England

INTRODUCTION

While there are regional childhood diabetes registers within the UK(1-6), across Europe(7) and worldwide(8), large scale national registers for childhood diabetes in Europe or elsewhere are not common outside Scandinavia. Following the St Vincent Declaration for Quality Assurance in Diabetes Care in 1989(9) several countries established diabetes registers to enable the quality audit of care(10, 11). In 2003 a National Diabetes Audit was established in England and Wales to assess the quality of care for diabetics, however this audit covers only 50% of children in England and Wales with diabetes(12).

With no national register of childhood diabetes, it is not possible to explore spatial and temporal trends, or assess the impact of potentially important environmental factors on diabetes incidence across England. Routinely collected morbidity data have been used previously to assess prevalence or incidence of diabetes in children(4, 13-17). Estimates of ascertainment of childhood diabetes cases from hospital admissions data have been reported to be 95% in the northern region of England over the years 1977-1986(13), 94% in Scotland over the years 1984-1994(4), and 86% in Northern Ireland over the years 1989-1994(15); however other papers report far lower ascertainment; 58% in the northern region and 74% in the south western region of England in 1988(14). Several limitations and issues associated with the use of routine data for this purpose have been highlighted. Problems of under-ascertainment of diabetes have been noted in some studies that have compared hospital records with other data sources(14, 17), however other studies have suggested problems of over-ascertainment(13, 16, 18). In addition, there may be substantial geographical differences in the way that childhood diabetes is managed across England. For example Leicestershire strives to manage diabetics as outpatients by offering treatment at auxiliary centres(19), in contrast Leeds Health Authority refers all newly diagnosed children, irrespective of diabetes type, to a hospital clinic(20).

These problems of under- and over-ascertainment, and the impact of regional management practices on ascertainment require exploration. Hospital records in England are known to have improved in quality since the introduction of National Health Service (NHS) reforms in April

1991(21, 22), and an updated assessment of the validity of this data source in England is long overdue. Our aim was to assess the feasibility of using hospital admissions data as a surrogate for a childhood diabetes register across England for subsequent use in epidemiological studies investigating the distribution and determinants of childhood diabetes.

METHODS

Data

The Small Area Health Statistics Unit (SAHSU) holds Hospital Episodes Statistics (HES) data from England for the period April 1992 to March 2006. All hospital admissions with a diabetes diagnosis (International Classification of Disease (ICD) revision 9 codes ‘250’ and ICD10 codes between ‘E10’ and ‘E14’) and where the patient age was in the range 0-14 years were extracted from the SAHSU HES dataset. Cases admitted to hospital with a primary diagnosis relating to type 2 diabetes (ICD9 category 250 codes with a fifth-digit sub classification ‘0’ or ‘2’ and ICD10 E11 codes) were excluded. The impact of the inclusion or exclusion of patients with an unspecified diabetes code was assessed. The dataset was restricted to patients aged 0-14 years, as the centralisation of delivery of clinical care in paediatric centres facilitates the completeness of ascertainment(23).

Data cleaning

The chronic nature of diabetes means many cases will be re-admitted, and thorough data cleaning was necessary to try to construct an ‘incident’ rather than ‘prevalent’ dataset i.e. to achieve a dataset where each child is counted only once. Duplicate records were identified on the basis of a shared date of birth, postcode and sex, and only the earliest admission was retained. The use of date of birth, postcode and sex as an approximate unique patient identifier is believed to confer at least 90% accuracy (i.e. remove at least 90% of the duplicate records) (21).

Unique patient identifiers do exist in the HES data, for example the HESID which *“uniquely identifies a patient across all data years…generated by matching records for the same patient using a combination of NHS Number, local patient identifier, postcode, sex and date of birth”* (24). This field, introduced in 2000, will not identify re-admissions throughout the whole data series, nonetheless, the completeness and potential usefulness of this field from 2000 onwards has been explored.

Data validation

The cleaned hospital admissions data were validated against regional population based diabetes register data, available from the Yorkshire Register of Diabetes in Children and Young People and the Bart's-Oxford (BOX) study of childhood diabetes. The Yorkshire register identifies children with insulin dependent diabetes mellitus resident in the counties of West Yorkshire, North Yorkshire and Humberside, and has data from 1978 to the current time, with an estimated ascertainment of more than 99%(23). The BOX study covers the former Oxford Regional Health Authority (Oxfordshire, Berkshire, Buckinghamshire and Northamptonshire and the Unitary Authorities that fall within their boundaries), with data available from 1986 to 2004, and an estimated ascertainment of 96%(3). Records pertaining to cases of childhood type 1 and type unknown diabetes in areas covered by each register were identified from the hospital admissions dataset to allow a direct comparison of case counts, by sex, age-group and year, to be made with case counts from each register, using Pearson correlation co-efficients. This validation aimed to highlight the degree of over- or under-ascertainment of the hospital admissions derived data for these regions, and enabled an exploration of the quality of the data over time (e.g. to establish whether the earlier years of data are of a high enough quality to be used). As well as assessing the validity of the HES data for all years/part years for which there was an overlap with the register data, the match was also assessed for the periods Jan 1995 – Dec 1999 and Jan 2000 – March 2006 (Jan 2000 – Dec 2004 for the BOX study) to reflect improvements in data quality over time, including the move to ICD10 in April 1995, and the introduction of the HESID in April 2000 which aided the removal of re-admissions.

A more detailed validation was also carried out to assess the degree of overlap of cases (i.e. to check the same cases are being identified by each data source) using data from the Yorkshire register for the period April 1992 to March 2006. This validation process involved the exchange of pseudonymised hospital admissions and Yorkshire register data, and independent data matching of the encrypted data. HES and Yorkshire register records were pseudonymised using the same 256-bit secure hash SHA256 algorithms to encrypt the personally identifiable fields (postcode, NHS number, date of birth); matching was undertaken on the basis of encrypted postcode, NHS number, date of birth, as well as admission year and sex. This exercise aimed to identify: i) cases that appear in both datasets, ii) cases that appear only in the admissions data, and iii) cases that appear only in the Yorkshire register. For cases that appeared only in the hospital admissions derived data (subset ii) or Yorkshire registry data (subset iii), the SAHSU and Yorkshire register teams respectively explored potential spatial, temporal, age and sex characteristics of these subsets to explain the discrepancies observed and provide a better understanding of any over- or under-ascertainment in the hospital admissions derived dataset.

Exploration of basic trends in childhood diabetes

Basic descriptive outputs were generated to explore the age distribution of diabetes by sex, the seasonality of onset, and the incidence of diabetes over time (using denominator data, available by sex and five year age group, from the 1991 and 2001 censuses, with changes in the population during the intervening years being estimated using mid-year population estimates and linear interpolation methods)(25).

Ethics

SAHSU has ethical approval from the Imperial College Research Ethics Committee for studies using routine health data. The National Information Governance Board for Health and Social Care (NIGB) approved the transfer of the pseudonymised data between SAHSU and the Yorkshire register for this study.

RESULTS

Data and data cleaning

From the extracted HES data (n=66,184 admissions), 58,230 (88%) of the diabetes admissions referred to type 1 diabetes, 3,774 (5%) to type 2 diabetes (which were therefore excluded), and 4,680 (7%) of admissions did not specify the diabetes type. For those admissions coded using ICD9 (between April 1992 and March 1995, n=10,750) these percentages were 63%, 21% and 16% respectively; for admissions coded using ICD10 (from April 1995 to March 2006, n=55,434) these percentages were 93%, 2% and 5% respectively.

When data were cleaned to retain only the first admission for each case with a unique date of birth, postcode and sex, the dataset contained 31,582 cases of type 1 diabetes (54.2% of the pre-cleaned type 1 data) and 2,149 cases of type unknown diabetes (45.9% of the pre-cleaned type unknown data).

HESIDs were available for all records from April 2000, so removal of duplicate cases on the basis of shared HESID only affected cases admitted from April 2000 onwards. Removal of admissions with duplicate HESIDs resulted in a type 1 diabetes dataset of 30,550 cases, and type 1 + type unknown dataset of 32,665 cases.

The ‘type 1’ and ‘type 1 + type unknown’ datasets were similar, and increasingly so after April 1995, corresponding to the move to ICD10 which resulted in a decrease in the number of ‘unknown’ diagnoses (from 16% under ICD9 to 5% under ICD10). The complete dataset (type 1 + type unknown) case counts were a better match overall and correlated better with the Yorkshire register and BOX family study data (e.g. the Pearson correlation co-efficients with the register data were 0.86, $p<0.01$ and 0.89, $p<0.01$ for Yorkshire, and 0.82, $p<0.01$ and 0.84, $p<0.01$ for the BOX family study, for the type 1 and type 1 + type unknown datasets respectively), and as such a pragmatic decision was taken to use this type 1 + type unknown dataset in subsequent analyses, and it is output from this dataset that is presented throughout this paper.

Data validation

Case counts by year, sex and age group for the relevant geographic areas from the hospital admission derived data were compared with case counts from the Yorkshire register and BOX family study.

For the Yorkshire area, over the period April 1992 to March 2006, the Pearson correlation co-efficient between annual case counts by sex and age-group for the hospital admissions and register data was 0.89, $p<0.01$, with the hospital admissions data explaining 79.7% of the variability in the Yorkshire register data (figure 1a and table 1). As can be seen from figure 1c and e, and table 1, the correlation coefficients increased for the periods January 1995 to Dec 1999 and January 2000 to March 2006. Correlations by sex were similar (data not shown), but were higher for the 0-4 and 5-9 age groups than the 10-14 age group, and again increased over time (table 1). The hospital admissions data tended to under-estimate case numbers in the younger age groups, and over-estimate case numbers in the older age groups (figure 1a, c and e).

For the BOX study, over the period January 1993 to December 2004, the Pearson correlation co-efficient between annual case counts by sex and age-group for the hospital admissions and register data was 0.84, $p<0.01$, with the hospital admissions data explaining 70.9% of the variability in the BOX study data (figure 1b and table 1). As can be seen from figure 1d and f, and table 1, correlation coefficients increased from the period January 1995 - December 1999 to January 2000 to December 2004. Correlations by sex were similar (data not shown). As with the Yorkshire area, the hospital admissions derived data tended to over-estimate case numbers in the older age groups (figure 1b, d and f).

(Table 1 and Figure 1 here)

There were some apparent anomalies in the data, for example case counts from the hospital admissions derived data for the year 1994 appeared artificially low for males and females of all age groups compared to the Yorkshire data (data not shown).

Cases for the Yorkshire area over period April 1992 – March 2006 from both the hospital admissions dataset and Yorkshire register were pseudonymised and independently matched using postcode, NHS number, date of birth, admission year and sex. Over this period a total of 2,236 cases aged 0-14 were recorded in the hospital admissions dataset and 2,312 recorded in the Yorkshire registry. Of the 2,236 cases in the hospital admissions dataset, 1,614 (72.2%) could be matched to cases in the Yorkshire register. Of the 2,312 cases in the Yorkshire register, 1,607 (69.5%) cases aged 0-14 were matched to patients in the hospital admissions dataset, and a further 7 (0.3%) aged over age 14 were also matched. The matching improved slightly when the diagnosis date of Yorkshire register records was not restricted to the period April 1992 – March 2006 (1,695/2,236 (76%) records in the Yorkshire register were matched to HES data). The availability of HESIDs from April 2000 improved the match considerably, with 1093/1204 (90.8%) of the hospital admission cases matching cases in the Yorkshire register, over the period 2000-2006, suggesting that duplicate cases in the HES data were a problem in the period prior to April 2000 before the introduction of the HESID. When looking at cases aged 0-9 years over the whole study period 1002/1159 (86.5%) of the hospital admission cases matched cases in the Yorkshire register.

Exploration of basic trends in childhood diabetes

The age distribution of diabetes by sex showed plateaus at 3-5 years (males and females), and peaks at 11-12 years (females) and 13 years (males) (figure 2).

(Figure 2 here)

The data also showed some seasonality in admissions in males and females, especially in the older male age groups (figure 3).

(Figure 3 here)

Overall incidence was 24.99 (95% Confidence Interval (CI) 24.71-25.26) per 100,000, and was similar for males (25.23 (95% CI 24.85-25.62) and females 24.73 (95% CI 24.34-25.11) per 100,000. Incidence increased by age group, from 14.31 (95% CI 13.95-14.67), 24.64 (95% CI 24.17-25.10) and 35.70 (95% CI 35.14-36.26) per 100,000 in males and females aged 0-4, 5-9 and 10-14 years respectively.

From 1995 to 2000 there was an increase in diabetes incidence in males and females of all age groups (figure 4), although after 2000 rates appeared to level off. From 1992 to 1994 the hospital admissions derived data show a downward trend in incidence, most pronounced in the 5-9 and 10-14 year old age groups.

(Figure 4 here)

DISCUSSION

There is currently no national childhood diabetes register for England. Such a register, covering the heterogeneous population of England and exposure differentials experienced across the country, is needed for exploring the causes of this important disease; type 1 diabetes was estimated to cost the UK approximately £212m in 2001 prices, with indirect costs likely to exceed this figure(26). We have shown that the HES data for England, which include all admissions to hospital for children diagnosed with diabetes (approximately 32,500 unique cases) over the period April 1992-March 2006, may serve as a proxy for a dedicated national diabetes register, until such a register is established.

Validation against available regional population based diabetes register data indicated that the hospital derived data improved in quality over time, and that data quality was better for 0-9 year old than for 10-14 year old children. As discussed below, basic descriptive outputs from the hospital derived data matched previously reported trends in age distribution of diabetes and seasonality of onset, and overall incidence matched well with previously reported findings for various regions of the UK(4, 7, 14, 15, 27-31).

However, there are several limitations to using HES data as a proxy for a dedicated diabetes register. Firstly, the HES data include only those children who developed diabetes and who were admitted to hospital; cases managed exclusively in the community setting will not be captured by these data, and our proxy register might, therefore, be biased towards those more severe cases unlikely to be stabilised as outpatients, and more likely to be re-admitted following discharge. Our incidence rates and validation exercises suggest that, overall, the HES data do capture a majority of cases, especially for younger cases and in the most recent years. However there are known regional differences in the management of childhood diabetes, with some areas dedicated to providing outpatient treatment(19, 32), in contrast to other areas where all newly diagnosed children are referred to a hospital clinic(20), meaning the hospital admissions data may better reflect incidence in some areas than others. One area in particular, Leicestershire, has reportedly attempted to manage diabetic children as outpatients rather than inpatients in hospital(19). The Leicestershire area maintained a population based childhood diabetes register up until March 1996 but it was not possible to obtain data from this register for validation of the HES dataset. Incidence rates for 0-9 year olds for the districts making up Leicestershire over the period 1995-1999 ranged from 0 to 11.86 (mean 7.45 (SD4.46)), and were significantly lower than rates in other districts ($p<0.01$), possibly reflecting this policy of outpatient treatment. However, incidence rates for 0-9 year olds over the period 2000-2005 ranged from 15.7 to 26.8 (mean 21.64 (SD 4.16)), and were not statistically significantly different to rates in other districts, suggesting that the more recent data effectively capture the anticipated number of cases in this area.

Secondly, the chronic nature of diabetes means that many children are re-admitted to hospital, and cleaning the hospital data to achieve an incident rather than prevalent dataset is a challenge, particularly given the lack of unique patient identifiers over the whole period of interest, and possibility of cases being diagnosed before our data series commencement. Again, our incidence rates and validation exercises suggest that, overall, the cleaned HES data do approximate an incident dataset, although approximately 5% (81/1695) of the hospital admissions cases matched to registry cases outside of the study period (i.e. cases likely to have first been admitted (diagnosed) prior to April 1992), and the admissions data tend to over-estimate case numbers in the older age group. The over-estimation of older cases likely reflects the fact that older children will have had more opportunity to move house between admissions, making more difficult the identification of duplicate case (i.e. re-admissions), particularly for cases first admitted prior to the introduction of the HESID.

Thirdly, the hospital admissions derived data cover a period of 14 years during which time there have been a number of changes that are likely to have improved data quality. These changes include a series of NHS reforms from 1991(21, 22), a move to ICD10 coding from April 1995, and the inclusion of an unique personal identifier (HESID) from April 2000. Validation of the hospital admissions data against available regional register data corroborate these improvements in the HES data over time. In addition to these acknowledged influences on data quality over time, there were also some apparent anomalies in the data, for example case counts from the hospital admissions derived data for the year 1994 appeared artificially low (figure 4). It seems likely that this anomaly relates, in part, to the move in HES from ICD9 to ICD10 in April 1995. Prior to this change-over, 21% of childhood diabetes admissions were classified as type 2 diabetes, but after the more to ICD10 this proportion dropped to only 2%. It seems very unlikely that the number of children suffering from type 2 diabetes differed to such a large extent over a period of just a few years. The National Diabetes Audit indicates that since 2003/2004 less than 2% of diabetes registrations in England and Wales were for type 2 diabetes(12), suggesting that some of these type 2 admissions before April 1995 may have been mis-diagnosed, and their exclusion from our dataset will have meant the exclusion of a significant number of type 1 and/or type unknown cases.

Overall, these findings suggest that HES data from April 2000 onwards, when disease coding used ICD10, and when HESID was available to aid in the removal of duplicate cases, is a good surrogate for a national register for children ages 0-9 years.

Basic trends in the age distribution, seasonality of onset, and incidence match well with previously reported findings. The age distribution of diabetes showed peaks at 3-5 years and at 11-12 years (females) and 13 years (males), as past work has shown(4, 14, 28, 29), and the data showed some seasonality in admissions, with more cases admitted during the winter months, especially in the older age groups, similar to previous reports(4, 7, 15, 27-29, 31). It has been suggested that these age and seasonal peaks are linked to the time of starting and changing schools(28), infections(29, 33) or pubertal changes(29), although none of these explanations have been found to be entirely satisfactory. The overall incidence rate of 24.99 per 100,000 was in line with published figures for the UK over the periods 1994-1998 and 1999-2003 of 24.7/29.8 (Northern Ireland), 21.7/22.4 (Oxford) and 19.7/23.3 (Yorkshire) (34). Incidence increased by age group, with males and females aged 5-9 having 1.61 and 1.85 times higher incidence than the 0-4 year age group, roughly in line with pooled data from around the world (1.52 and 1.72 times the incidence for males and females respectively) (30). However, the 10-14 year old age groups had a higher than anticipated incidence compared to the 0-4 year old age group (2.42 and 2.58 times the incidence versus worldwide figures of 1.94 and 1.93 in males and females respectively) (30). This higher ratio of incidence in the 10-14 years olds could be a reflection of the over-ascertainment in this age group, and/or an indication that worldwide figures do not accurately represent the situation in England, although data from the Yorkshire register over the period 1978-1998 do support these worldwide ratios(35). The increasing trend in diabetes incidence observed in previous studies worldwide(8, 30), from Europe(7, 36, 37) and from the UK(3, 4, 23) was also evident in the hospital admissions data over the period 1995 to 2000. Nonetheless there were inconsistencies between time trends observed in the HES data and those observed elsewhere for the period pre 1995, when the HES data showed a decrease in incidence in the 5-9 and 10-14 year old age groups. Data from Scotland for the years 1984 to 1993 showed a 2% increase in incidence per year(4); data from the Oxford region over the years 1985 to 1996 showed a 4% increase in incidence per year(3); and data from Yorkshire over the years 1978 to 2000 showed a 2.9% increase in incidence per year(23), with none of these studies suggesting a decrease in incidence in the years 1992-1994. Again, this discrepancy is likely to be due, in part, to the move in HES from ICD9 to ICD10 in April 1995, and the impact this change-over had on the proportion of admissions classified as type 2 diabetes. As discussed above, it seems probable that some of these admissions for type 2 diabetes up to the end of March 1995 may have been mis-diagnosed, and by excluding these cases from our dataset we may have under estimated the number of type 1 and/or type unknown cases. After 2000, incidence in the HES data appears to level off, or even decrease in the 10-14 years old female age group. One report suggested that the rise in incidence in some high risk regions (e.g. northern Europe) may have reached a plateau(7), however others do not support this conclusion(30, 34).

In summary, in lieu of a national register, HES data for more recent years (after April 2000), and for younger children (aged 0-9 years), provide a useful surrogate with which to explore spatial and temporal trends in data quality, diabetes management and diabetes incidence across England. Further exploration of this data resource will highlight the value of establishing a dedicated national diabetes register, and/or incorporating additional data sources in this surrogate (including laboratory, retinal screening, and Primary Care data) for undertaking aetiological research into this important childhood condition.

Acknowledgements

We thank the Hospital Episode Statistics (The NHS Information Centre for health and social care) for provision of hospital admissions data, and the United Kingdom Office for National Statistics for provision of population and socio-economic data.

The work of the Small Area Health Statistics Unit was funded by a grant from the Department of Health for England and the United Kingdom Department for Environment, Food and Rural Affairs. The SAHSU steering committee approved of the study being undertaken as part of the SAHSU work programme, but had no role in the study design, collection, analysis, or interpretation of the data, in the writing of the report, or in the decision to submit the article for publication.

The Yorkshire Register of Diabetes in Children and Young People is funded by the Healthcare Quality Improvement Partnership.

We would like to acknowledge the Bart's-Oxford (BOX) Study Group, Professor Polly Bingley and Isabel Wilson for providing data from this study for use in the validation of the HES extract. The BOX study was funded by Diabetes UK.

REFERENCES

1. Burden AC, Hearnshaw JR, Swift PG. Childhood diabetes mellitus: an increasing incidence. Diabetic Medicine 1989; 6(4):334-6.

2. Cardwell CR, Carson DJ, Patterson CC. Higher incidence of childhood-onset type 1 diabetes mellitus in remote areas: a UK regional small-area analysis. Diabetologia 2006; 49:2074-2077.

3. Gardner SG, Bingley PJ, Sawtell PA, Weeks S, Gale EA. Rising incidence of insulin dependent diabetes in children aged under 5 years in the Oxford region: time trend analysis. The Bart's-Oxford Study Group. BMJ 1997; 315(7110):713-7.

4. Rangasami JJ, Greenwood DC, McSporran B, Smail PJ, Patterson CC, Waugh NR. Rising incidence of type 1 diabetes in Scottish children, 1984-93. The Scottish Study Group for the Care of Young Diabetics. Archives of Disease in Childhood 1997; 77(3):210-3.

5. Staines A, Bodansky HJ, Lilley HE, Stephenson C, McNally RJ, Cartwright RA. The epidemiology of diabetes mellitus in the United Kingdom: the Yorkshire Regional Childhood Diabetes Register. Diabetologia 1993; 36(12):1282-7.

6. Zhao HX, Moyeed RA, Stenhouse EA, Demaine AG, Millward BA. Space-time clustering of childhood Type 1 diabetes in Devon and Cornwall, England. Diabetic Medicine 2002; 19(8):667-672.

7. Green A, Patterson CC, Group ETS. Trends in the incidence of childhood-onset diabetes in Europe 1989-1998. Diabetologia 2001; 44 Suppl 3:B3-B8.

8. Karvonen M, Viik-Kajander M, Moltchanova E, Libman I, LaPorte R, Tuomilehto J. Incidence of childhood type 1 diabetes worldwide. Diabetes Mondiale (DiaMond) Project Group. Diabetes Care 2000; 23(10):1516-1526.

9. Anonymous. Diabetes care and research in Europe: the Saint Vincent declaration. Diabetic Medicine 1990; 7(4):360.

10. Carstensen B, Kristensen JK, Ottosen P, Borch-Johnsen K, Steering Group of the National Diabetes R. The Danish National Diabetes Register: trends in incidence, prevalence and mortality. Diabetologia 2008; 51(12):2187-96.

11. Gudbjornsdottir S, Cederholm J, Nilsson PM, Eliasson B, Steering Committee of the Swedish National Diabetes R. The National Diabetes Register in Sweden: an implementation of the St. Vincent Declaration for Quality Improvement in Diabetes Care. Diabetes Care 2003; 26(4):1270-6.

12. National Diabetes Audit. Report for the audit period 2007-2008, NHS The Information Centre. 2008.

13. Crow YJ, Alberti KG, Parkin JM. Insulin dependent diabetes in childhood and material deprivation in northern England, 1977-86. BMJ 1991; 303(6795):158-60.

14. Metcalfe MA, Baum JD. Incidence of insulin dependent diabetes in children aged under 15 years in the British Isles during 1988. BMJ 1991; 302(6774):443-7.

15. Patterson CC, Carson DJ, Hadden DR. Epidemiology of childhood IDDM in Northern Ireland 1989-1994: low incidence in areas with highest population density and most household crowding. Northern Ireland Diabetes Study Group. Diabetologia 1996; 39(9):1063-9.

16. Waugh NR. Insulin-dependent diabetes in a Scottish region: incidence and urban/rural differences. Journal of Epidemiology & Community Health 1986; 40(3):240-3.

17. Williams DR, Fuller JH, Stevens LK. Validity of routinely collected hospital admissions data on diabetes. Diabetic Medicine 1989; 6(4):320-4.

18. Patterson CC, Waugh NR. Urban/rural and deprivational differences in incidence and clustering of childhood diabetes in Scotland. International Journal of Epidemiology 1992; 21(1):108-17.

19. Swift PG, Hearnshaw JR, Botha JL, Wright G, Raymond NT, Jamieson KF. A decade of diabetes: keeping children out of hospital. BMJ 1993; 307(6896):96-8.

20. Feltbower RG, McKinney PA, Campbell FM, Stephenson CR, Bodansky HJ. Type 2 and other forms of diabetes in 0-30 year olds: a hospital based study in Leeds, UK. Archives of Disease in Childhood 2003; 88(8):676-9.

21. Hansell A, Bottle A, Shurlock L, Aylin P. Accessing and using hospital activity data. Journal of Public Health Medicine 2001; 23(1):51-6.

22. Hyndman SJ, Williams DR, Merrill SL, Lipscombe JM, Palmer CR. Rates of admission to hospital for asthma. BMJ 1994; 308(6944):1596-600.

23. Feltbower RG, McKinney PA, Parslow RC, Stephenson CR, Bodansky HJ. Type 1 diabetes in Yorkshire, UK: time trends in 0-14 and 15-29-year-olds, age at onset and age-period-cohort modelling. Diabetic Medicine 2003; 20(6):437-441.

24. HES online. 2009 Available from: http://www.hesonline.nhs.uk/Ease/servlet/ContentServer?siteID=1937&categoryID=571&dirID=110579&headerID=.

25. Arnold R, Elliott P, Wakefield J, Quinn M. Population counts in small areas: implications for studies of environment and health. 1997, Office for National Statistics, Government Statistical Service.

26. Royal College of Physicians. Type 1 diabetes in adults. National clinical guideline for diagnosis and management in primary and secondary care. 2004, National Collaborating Centre for Chronic Conditions.

27. Patterson CC, Smith PG, Webb J, Heasman MA, Mann JI. Geographical variation in the incidence of diabetes mellitus in Scottish children during the period 1977-1983. Diabetic Medicine 1988; 5(2):160-5.

28. Patterson CC, Thorogood M, Smith PG, Heasman MA, Clarke JA, Mann JI. Epidemiology of type 1 (insulin-dependent) diabetes in Scotland 1968-1976: evidence of an increasing incidence. Diabetologia 1983; 24(4):238-43.

29. Bloom A, Hayes TM, Gamble DR. Register of newly diagnosed diabetic children. British Medical Journal 1975; 3(5983):580-3.

30. Diamond Project Group. Incidence and trends of childhood Type 1 diabetes worldwide 1990-1999. Diabetic Medicine 2006; 23(8):857-66.

31. Green A. Descriptive epidemiology of type 1 diabetes in youth: incidence, mortality, prevalence, and secular trends. Endocrine Research 2008; 33(1-2):1-15.

32. Lowes L, Gregory JW. Management of newly diagnosed diabetes: home or hospital? Archives of Disease in Childhood 2004; 89(10):934-7.

33. Hyoty H, Taylor KW. The role of viruses in human diabetes. Diabetologia 2002; 45(10):1353-61.

34. Patterson CC, Dahlquist GG, Gyurus E, Green A, Soltesz G, Group ES. Incidence trends for childhood type 1 diabetes in Europe during 1989-2003 and predicted new cases 2005-20: a multicentre prospective registration study. Lancet 2009; 373(9680):2027-33.

35. Feltbower RG, McKinney PA, Bodansky HJ. Rising incidence of childhood diabetes is seen at all ages and in urban and rural settings in Yorkshire, United Kingdom. Diabetologia 2000; 43(5):682-684.

36. Eurodiab Ace Study Group. Variation and trends in incidence of childhood diabetes in Europe. EURODIAB ACE Study Group. Lancet 2000; 355(9207):873-876.

37. Green A, Gale EA, Patterson CC. Incidence of childhood-onset insulin-dependent diabetes mellitus: the EURODIAB ACE Study. Lancet 1992; 339(8798):905-909.

TABLES

Table 1: Case counts by age-group and period for the hospital admissions derived data (HES) and the Yorkshire Register and Bart's-Oxford (BOX) family study data (Register), with Pearson correlation coefficients between HES and Register case counts, by sex, age-group and year.

	Case count HES/Register	Correlation coefficient (p-value)	95% CI	Case count HES/Register	Correlation coefficient (p-value)	95% CI	Case count HES/Register	Correlation coefficient (p-value)	95% CI
Yorkshire	**Apr 1992 - Mar 2006**			**Jan 1995 – Dec 1999**			**Jan 2000 - Mar 2006**		
All ages	2224/2161	0.89 (<0.01)	0.84-0.93	782/759	0.93 (<0.01)	0.86-0.97	1204/1033	0.96 (<0.01)	0.93-0.98
0 – 4 years	401/442	0.84 (<0.01)	0.69-0.92	149/167	0.86 (<0.01)	0.50-0.97	212/203	0.98 (<0.01)	0.94-0.99
5 – 9 years	749/763	0.84 (<0.01)	0.69-0.92	271/271	0.87 (<0.01)	0.53-0.97	396/363	0.96 (<0.01)	0.88-0.99
10 – 14 years	1074/956	0.79 (<0.01)	0.60-0.90	362/321	0.87 (<0.01)	0.53-0.97	596/467	0.91 (<0.01)	0.73-0.97
BOX study	**Jan 1993 – Dec 2004**			**Jan 1995 – Dec 1999**			**Jan 2000 - Dec 2004**		
All ages	1465/1435	0.84 (<0.01)	0.76-0.90	577/600	0.77 (<0.01)	0.57-0.89	731/628	0.96 (<0.01)	0.92-0.98
0 – 4 years	257/310	0.39 (0.06)	-0.02-0.69	88/127	0.64 (0.05)	0.02-0.91	130/125	0.74 (0.02)	0.21-0.93
5 – 9 years	473/472	0.76 (<0.01)	0.51-0.89	210/232	0.59 (0.07)	-0.06-0.89	213/190	0.87 (<0.01)	0.53-0.97
10 – 14 years	735/653	0.67 (<0.01)	0.37-0.85	279/241	0.36 (0.30)	-0.35-0.81	388/313	0.77 (0.01)	0.27-0.94

Figure 1: Correlation between the hospital admissions and register data plotted as annual case counts by age group and sex, for a) Yorkshire, April 1992 to March 2006, b) former Oxford health region, Jan 1993 – Dec 2004, c) Yorkshire, Jan 1995 to Dec 1999, d) former Oxford health region, Jan 1995 – Dec 1999, e) Yorkshire, Jan 2000 to March 2006, f) former Oxford health region, Jan 2000 – Dec 2004

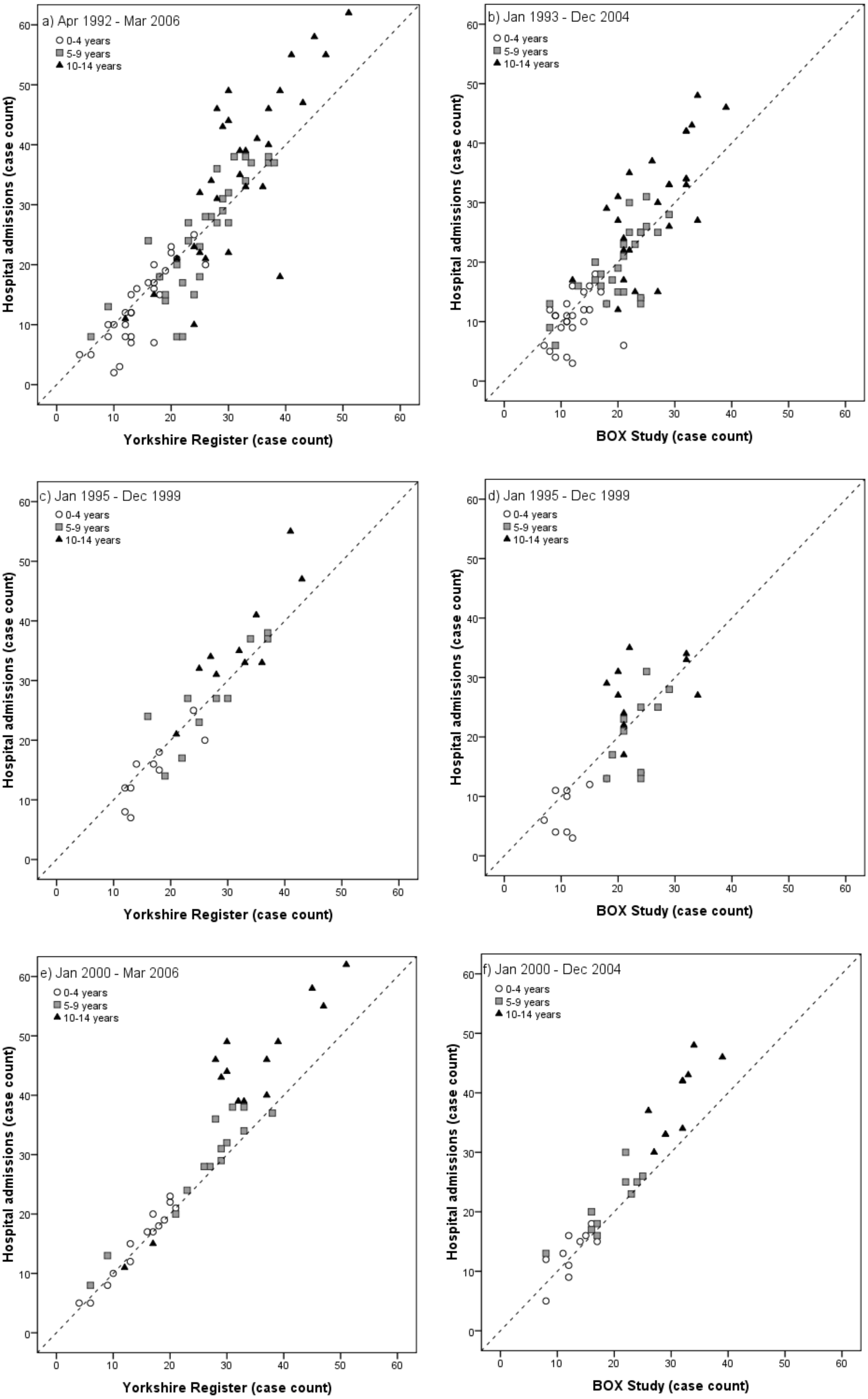
a) Apr 1992 - Mar 2006
0-4 years
5-9 years
10-14 years
Hospital admissions (case count)
Yorkshire Register (case count)
b) Jan 1993 - Dec 2004
0-4 years
5-9 years
10-14 years
Hospital admissions (case count)
BOX Study (case count)
c) Jan 1995 - Dec 1999
0-4 years
5-9 years
10-14 years
Hospital admissions (case count)
Yorkshire Register (case count)
d) Jan 1995 - Dec 1999
0-4 years
5-9 years
10-14 years
Hospital admissions (case count)
BOX Study (case count)
e) Jan 2000 - Mar 2006
0-4 years
5-9 years
10-14 years
Hospital admissions (case count)
Yorkshire Register (case count)
f) Jan 2000 - Dec 2004
0-4 years
5-9 years
10-14 years
Hospital admissions (case count)
BOX Study (case count)

Figure 2: Age distribution of diabetes cases by sex, over the period April 1992 to March 2006

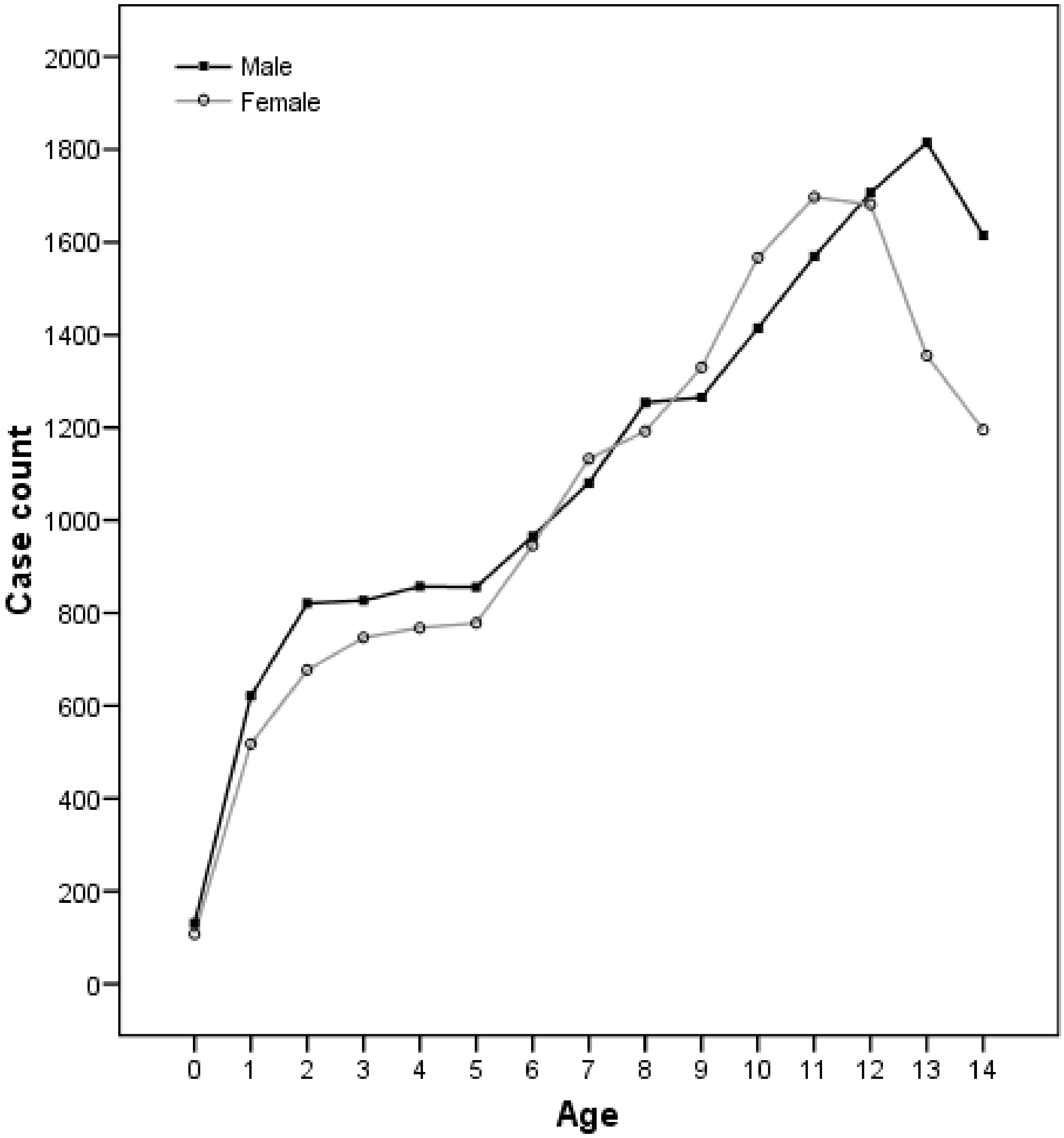

Figure 3: Seasonality of admissions by age group for a) males and b) females, over the period April 1992 to March 2006

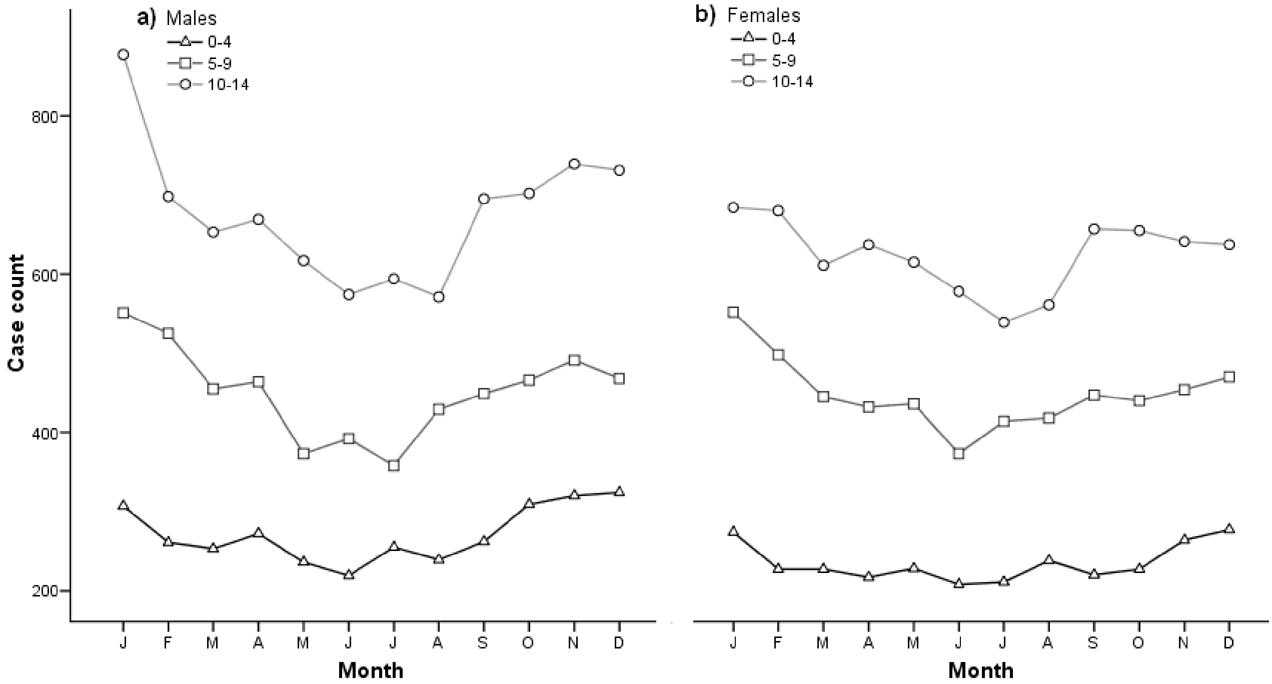

Figure 4: Time trends in age-specific diabetes incidence rates for a) males and b) females, over the period April 1992 to December 2005