Type 1 diabetes prevalence increasing globally and regionally: the role of natural selection and life expectancy at birth

Wen-Peng You,1 Maciej Henneberg1,2

ABSTRACT

Objective: Prevalence of type 1 diabetes (T1D) disease is increasing worldwide. We aim to test correlation of T1D prevalence to the reduced natural selection measured by Biological State Index (Ibs).

Research design and methods: Country-specific estimates of T1D prevalence, life expectancy, obesity prevalence rate, urbanization rates, per capita sugars consumption and per capita gross domestic product (GDP) were obtained. Ibs and country-specific longevity (e50) increase for each country were self-calculated. These data were then matched to T1D prevalence by country for our ecological study among 118 countries. Countries were also grouped to study the associations in different regions. SPSS V.22 was used for correlation analysis.

Results: Worldwide, both Ibs and life expectancy at birth (Ibs proxy) were significantly correlated to T1D prevalence in Pearson r (r=0.713, p<0.001 and r=0.722, p<0.001, respectively) and Spearman’s r (r=0.724, p<0.001 and r=0.689, p<0.001, respectively). T1D prevalence was not correlated to longevity increase measured as life expectancy at 50 years old. T1D prevalence was significantly associated with Ibs (r=0.307, p<0.001) and newborn life expectancy (r=0.349, p<0.001) independent of per capita total sugar consumption, per capita GDP, urbanization and obesity prevalence in partial correlation. Globally, both life expectancy at birth and Ibs exponentially correlated to T1D prevalence. Pearson correlations generally existed in different country categorizations by geographic region, culture background and economic status.

Conclusions: Reduced natural selection may have contributed to the increasing T1D prevalence worldwide. T1D epidemiology study in total population may be the practical solution to identify the causes of increasing T1D prevalence.

BACKGROUND

Type 1 diabetes (T1D) is an autoimmune disease with a strong genetic component.1 2 It can occur at any age, but tends to develop in childhood,3 so it has long been called ‘juvenile diabetes’. T1D is characterized by destruction of pancreatic β-cells, culminating in absolute insulin deficiency.4 As of 2014, an estimated 387 million people have diabetes worldwide,5 of which T1D accounts for between 5% and 10%.6 Diabetic complications continue to be a major cause of morbidity and mortality in persons with T1D.7 Great efforts have been made to assess the incidence and prevalence of T1D. Unfortunately, the exact etiology and pathogenesis of T1D is still unknown. Generally, longitudinal or cross-sectional studies are often locally or regionally performed. Consequently, it is difficult to access generalizable results because the epidemiology of T1D is known to be heterogeneous with regard to geography and ethnicity. Genetic predisposition to T1D is only alleged to explain some of the geographic variability in T1D occurrence, but it cannot account for its rapidly increasing frequency.8 A number of studies have associated gross domestic product (GDP) level with T1D prevalence or incidence,8-11 but GDP does not fully explain variations and trends in T1D prevalence rates observed in many countries, for example Japan. It has been postulated that environmental factors may be able to trigger an autoimmune destruction of the β-cells leading to absolute dependence on insulin treatment,8 9 12-17 however, these environmental factors are circumstantial.13 Natural selection, as one of the basic mechanisms of evolution, is the differential survival and fertility of individuals due to...
differences in phenotype that reflect genetic differences. In our modern society, natural selection still acts on all members of a population, selecting those individuals that have an increased reproductive success (survival and/or fertility).\textsuperscript{18} The ‘Biological State Index (I\textsubscript{bs})’ has been proposed to measure the populational reproductive success by taking into account potential loss of reproductive success by dying at age x, summed over all age categories.\textsuperscript{19,20} The I\textsubscript{bs} is calculated by combining age-specific death frequency (dx variable of a life table) with an age-specific reproductive loss (s\textsubscript{x}):

\[ I_{bs} = 1 - \sum_{x=0}^{x=100} dx \cdot Sx \]

where dx is the frequency of death at age x or represents the mortality rate, s\textsubscript{x} the reproductive loss from dying at age x, that is, the estimated probability of not possessing the complete number of births at age x. s\textsubscript{x} is based on the cumulative number of births at specific ages.\textsuperscript{20,21} The construction and interpretation of the I\textsubscript{bs} was predicated on the assumption that heritability of human fertility variance is negligible.\textsuperscript{22}

An I\textsubscript{bs} value of 1 indicates total adaptation of the population to their environment (ability to overcome selection pressures that are present). An I\textsubscript{bs} value of 0 signifies a total lack of adaptation (inability to overcome selection pressures that are present), and an impossibility to give life to the next generation. An I\textsubscript{bs} value close to 0 indicates large effective natural selection pressures acting on a population, since few individuals are surviving to produce offspring. In such a scenario there is a possibility for fast evolution, since many genes may not be passed to the next generation. An I\textsubscript{bs} value close to 1 indicates that natural selection is not having much effect on the population since many individuals are able to maximally contribute to producing the next generation. Thus, the I\textsubscript{bs} permits the estimation of the magnitude of the successful reproduction of a population.

The genetic trait of T1D may allow individuals from a population to pass their T1D genes on to their next generation. What fraction of a population had a chance to fully participate in reproducing under a given set of mortality conditions may be associated with the proportion of population carrying T1D genes in the next generation. Previously, Stephan and Henneberg\textsuperscript{23} raised a concern that the developed populations may have accumulated more unfavorable genes, such as T1D genes because natural selection (measured by the I\textsubscript{bs}) has been greatly reduced. Recently, a systematic review concluded that the T1D prevalence rate was associated with age increase in population.\textsuperscript{3} Therefore, the objective of the current paper was to use country-specific data to test, from a global perspective, that population with greater I\textsubscript{bs} value, fuelled by life expectancy at birth, may have higher T1D prevalence, using empirical, macro-level data collected from the major international organizations.

**MATERIAL AND METHOD**

The dependent variable in the analysis was the country-specific estimate of T1D prevalence (sum of rates in 0–14 and 15-years old groups) which was published by the International Diabetes Federation (IDF) in 2000.\textsuperscript{24} We used I\textsubscript{bs} and life expectancy at birth of each country as the independent variables. The I\textsubscript{bs} was calculated as proposed by Henneberg\textsuperscript{20} and Henneberg and Piontek\textsuperscript{39} with the fertility data of each country published by United Nations in 2008\textsuperscript{25} and the mortality data of life tables (2009) published by WHO in 2012.\textsuperscript{26} Life expectancy at birth is a proxy for I\textsubscript{bs} since it expresses an opportunity of a newborn to survive to a specific age, which is usually an age falling into reproductive life span or above it. We extracted the country-specific life expectancy (years) at birth by country published by WHO in 2013.\textsuperscript{27} GDP,\textsuperscript{30-31} urbanization,\textsuperscript{8} and body weight status\textsuperscript{17} have been associated with T1D prevalence. It has been suggested that a population with greater sugar consumption may have greater diabetes (total) prevalence\textsuperscript{28} and that sugar consumption may affect health of patients with T1D.\textsuperscript{29} Therefore, we controlled for market availability of sugars and sweeteners (sugars in short) in g/capita/day in each country for year 2010 from United Nations Food and Agricultural Organization (FAO) Food Balance Sheet, per capita GDP (expressed in purchasing power parity in 2010 US dollars for comparability among countries) and urbanization (percentage of population living in urban areas in each country in 2010) from the World Bank’s World Development Indicators Database, and obesity prevalence rate (percentage of the population aged 18+ years old with body mass index (BMI) ≥30 kg/m\textsuperscript{2} in each country during 2010) from the WHO.

No ethical approval or written informed consent for participation was required in this study as all the aforementioned data were freely downloaded from the United Nations agents’ websites.

We aligned all independent variables and confounding factors with the country-specific T1D prevalence, and obtained a set of data for 118 countries. All country-specific data were put in a uniform format. Each country was treated individually and all of their available information was analysed. Data of calculated I\textsubscript{bs} and summary statistics are further described in the supporting information (see online supplementary tables S1 and S2).

In order to demonstrate the universal association between T1D prevalence and I\textsubscript{bs} and life expectancy at birth respectively in different country groupings, we categorized the countries for correlation analyses based on (1) the WHO regional classifications;\textsuperscript{30} (2) the strong contrast in terms of geographic distributions, per capita GDP levels and/or cultural backgrounds. We analysed the correlation in the six country groupings: Latin America and the Caribbean (LAC),\textsuperscript{33} the Arab World,\textsuperscript{32} the Organisation for Economic Co-operation and Development (OECD),\textsuperscript{33}...
European Economic Area (EEA), Asia Cooperation Dialogue (ACD) and the Asia-Pacific Economic Cooperation (APEC). In our analysis, we only included those countries for which we could access their data for the specific groupings.

To a large extent, grouping countries for analysis may also allow us to align our findings against previous local or regional studies regarding heterogeneous T1D epidemiology due to various geographic location and ethnicity.

It might be considered that T1D prevalence is a result of the increase in longevity rather than relaxed natural selection of the genetic background of T1D since general health and advances in medical care improve survival of patients with T1D. This consideration is clarified by much stronger correlation between T1D prevalence and Ibs than that between T1D and life expectancy (1950–2010). The rationale to use country-specific life expectancy (1950–2010) is that (1) the estimate of life expectancy based on this segment of the population may not be biased with child (0–15 years) mortality, in particular due to deaths caused by T1D disease, and (2) deaths of adults, especially females, during the reproductive life span (15–50 years) that may differentiate numbers of T1D genes passed on to new generations. Thus, we obtained country-specific life expectancies (e50) for the periods of 1950–1955 and 2005–2010. The rationale to use country-specific life expectancy (1950–2010) is that (1) the estimate of life expectancy based on this segment of the population may not be biased with child (0–15 years) mortality, in particular due to deaths caused by T1D disease, and (2) deaths of adults, especially females, during the reproductive life span (15–50 years) that may differentiate numbers of T1D genes passed on to new generations. Thus, we obtained country-specific life expectancies (e50) for the periods of 1950–1955 and 2005–2010 respectively from the WHO life tables. And then we calculated the life expectancy increase from the period of 1950–1955 to the period of 2005–2010 for each country producing a new variable, which is ‘life expectancy increase (e50, 1950–2010)’ for each country across the 55 years. We repeated the above correlation analysis after we replaced the variable of ‘life expectancy at birth’ with life expectancy increase (e50, 1995–2010).

Pearson’s correlation coefficient, Spearman’s r and partial correlation analyses were conducted using SPSS V.22 (SPSS Inc, Chicago Illinois, USA). In this study, the data were log transformed for correlation analysis in SPSS.

RESULTS
Worldwide T1D prevalence was in significant associations with Ibs using Pearson r (r=0.713, p<0.001) and Spearman’s r (r=0.724, p<0.001), respectively. The similar associations were also observed between T1D and life expectancy in Pearson model (r=0.722, p<0.001) and Spearman’s model (r=0.689, p<0.001), respectively (table 1). Further investigation with partial correlation analysis showed that worldwide the association between T1D prevalence was still strongly associated with Ibs (r=0.507, p<0.001) and life expectancy (r=0.349, p<0.001) when we controlled for per capita total sugars availability, per capita GDP, urbanization and obesity prevalence (table 1). All confounders were in significant associations with T1D prevalence rate in both Pearson r and Spearman’s r.

Globally, T1D is noted to be exponentially related with both Ibs (R2=0.5203) and life expectancy (R2=0.5302) (figures 1 and 2) after we removed two outliers of extremely high T1D prevalence in Finland and Sweden, respectively.

Worldwide, strong positive Pearson r coefficients were observed in the correlation of Ibs to life expectancy (proxy of Ibs, r=0.908, p<0.001) and to GDP (r=0.781, p<0.001).

Table 2 presented that Europe had the significant associations with Ibs (r=0.502, p=0.001) and life expectancy (r=0.610, p<0.001), respectively. The sub-Europe (EEA) also had the very strong associations of T1D prevalence worldwide.
to both I_{bs} (r=0.479, p=0.009) and life expectancy (r=0.574, p=0.001) (table 3). We observed the slightly negative correlations of T1D prevalence rate to the I_{bs} and life expectancy, respectively, in two Asian country groupings, SEARO (South-East Asia) and the ACD (table 3). Both I_{bs} and life expectancy were insignificantly associated with T1D prevalence rate in other four WHO regions, AFRO, AMRO, EMRO and WPRO (table 2).

Table 3 showed that the T1D prevalence rate was associated with both the I_{bs} and newborn life expectancy in the groupings consisting of countries with similar cultures (Arab World and EEA) and different cultures (APEC and OECD), with similar economy status (OECD) and those with economy status in disparity (APEC), and with the heterogeneous region (LAC) and the homogeneous area (EEA and ACD).

Table 4 showed that in general T1D prevalence is not correlated to longevity increase, but strongly correlated to I_{bs} at country level. Life expectancy increase (e50, 1950–2010) does not correlate significantly with T1D prevalence in Pearson correlation (r=0.165, p=0.079) or Spearman’s r (r=0.166, p=0.077). These contrasted with the correlation between I_{bs} and T1D prevalence in Pearson correlation (r=0.713, p<0.001) and in Spearman’s r (r=0.724, p<0.001), respectively. In partial correlation analysis, when we controlled for GDP, BMI≥30, urbanization and sugars intake, T1D was negatively and insignificantly correlated to life expectancy increase (e50, 1950–2010) (r=−0.070, p=0.487), but it was in strong and significant correlation to I_{bs} (r=0.276, p=0.005).

DISCUSSION

The worldwide trend of increased T1D prevalence likelihood has multiple etiologies, which may act through multiple mechanisms. By assessing the T1D prevalence rate data for 118 countries we have shown that globally and regionally population which had greater value of I_{bs} (less opportunity for natural selection) may have greater T1D prevalence and secondly, that newborn life expectancy was significantly associated with T1D prevalence rate at population level.

Overall, the operation of natural selection on contemporary populations is declining due to modern medicine, but the magnitude of the decline may differ between countries due to their specific level of sanitation, medical interventions and public health measures. Natural selection is still one of the major evolutionary forces that informs changes in gene frequencies in a population through the action of differential fertility and mortality over generations. For example, studies have shown that the increasing prevalence rates of a partially heritable disease, nasal septa and lacrimal bone defects may be attributed to the decreasing effect of natural selection. More than 40 genetic loci located in different chromosomes have been associated with T1D in multiple studies. Although T1D can be fatal, the majority of genetically predisposed people do not develop T1D. This allows for accumulation of genetic predisposition in human populations. This accumulation will increase when fewer persons who developed a disease would die. Differential fertility and mortality are the basic events of natural selection, which operate singly or jointly to determine the fitness (reproductive success) of a particular population in a given environment. The country-specific fertility-based and mortality-based I_{bs} at different levels may indicate their different successful reproduction opportunities of individuals in the succeeding countries. The reproduction success opportunity of each population may determine their magnitude of T1D genes accumulation, thus influenced prevalence rate of patients with T1D in their next generations. In the present study, the correlation of I_{bs}...
to the T1D prevalence rate has been observed, which was compatible with suggestion that lower opportunity for selection allows accumulation of unfavorable genes. Our analysis of correlations between T1D prevalence and $I_{bs}$ by region or by WHO grouped countries seems to indicate that in regions where insulin was available earlier and that had better availability of healthcare the relationship is stronger. This provides the analog of a snapshot what could happen at different times in the same region as time from insulin introduction and improvement of healthcare increased. Thus, the distribution across different populations could be interpreted as a surrogate measure of the evolution in time of T1D prevalence after the introduction of insulin. Artificial insulin introduced for T1D treatment and increasing insulin availability may have played a key role in reducing natural selection as insulin enables countless people with onset of T1D to survive and maintain normal reproductive capacity. This may have been boosting T1D genes accumulation and prevalence of T1D. T1D can affect people of any age, but usually occurs in children or sexually mature young adults who have greater potential to reproduce than older adults. T1D has been historically, and continues to be, the most common type of diabetes in children and adolescents. Insulin is the priority for T1D treatment. Otherwise, patient with T1D may only live up to 1 year, some only a few weeks. Several human generations have benefited from insulin since it was discovered and became available in early 1920s. Reduced natural selection boosted by insulin treatment of several generations may have enabled cumulative effect of T1D genes frequency in human population to occur quickly and to be noticeable for a couple of decades. Studies have shown that a partially heritable disease, phenylketonuria was only noticeable after being accumulated for several generations with about 2% increase in each generation.

T1D prevalence/incidence is increasing worldwide with special regard to the developed countries. This may be partially attributable to earlier and greater affordability of insulin, in addition to relatively more reduced natural selection (greater $I_{bs}$ values) in those developed countries. Although exogenous insulin can be obtained from animals (bovine and porcine), production, transportation, storage and administration of such insulin was extremely expensive, which may be beyond the affordability of many patients with T1D, especially those from developing countries. Biosynthetic insulin based on DNA technology has been commercially available since 1982 and it has been thought that it can continue to accommodate global demand because of low cost from the production to administration. However, unfortunately life-saving insulin is still less accessible, affordable, or both to people diagnosed with diabetes in a developing country than their counterparts in the developed world. This lower survivorship of patients with T1D may contribute to lower prevalence figures directly, besides the fact that less predisposing genes have accumulated in the gene pools of those countries.

### Table 3

<table>
<thead>
<tr>
<th>Country grouping</th>
<th>$I_{bs}$ Pearson’s r</th>
<th>Significance</th>
<th>Life expectancy at birth Pearson’s r</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>APEC (n=16)</td>
<td>0.340</td>
<td>0.197</td>
<td>0.369</td>
<td>0.160</td>
</tr>
<tr>
<td>Arab World (n=13)</td>
<td>0.748</td>
<td>0.003</td>
<td>0.469</td>
<td>0.106</td>
</tr>
<tr>
<td>EEA (n=29)</td>
<td>0.479</td>
<td>0.009</td>
<td>0.574</td>
<td>0.001</td>
</tr>
<tr>
<td>ACD (n=20)</td>
<td>–0.288</td>
<td>0.253</td>
<td>–0.392</td>
<td>0.087</td>
</tr>
<tr>
<td>OECD (n=34)</td>
<td>0.365</td>
<td>0.034</td>
<td>0.155</td>
<td>0.382</td>
</tr>
<tr>
<td>LAC (n=26)</td>
<td>0.524</td>
<td>0.006</td>
<td>0.044</td>
<td>0.831</td>
</tr>
</tbody>
</table>

ACD, Asia Cooperation Dialogue; APEC, Asia-Pacific Economic Cooperation; EEA, European Economic Area; $I_{bs}$, Biological State Index; LAC, Latin America and the Caribbean; OECD, Organisation for Economic Co-operation and Development; T1D, type 1 diabetes.

### Table 4

<table>
<thead>
<tr>
<th></th>
<th>Pearson’s r</th>
<th>Spearman’s r</th>
<th>Partial correlation*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r</td>
<td>p Value</td>
<td>n</td>
</tr>
<tr>
<td>Log $I_{bs}$</td>
<td>0.713</td>
<td>&lt;0.001</td>
<td>118</td>
</tr>
<tr>
<td>Log life expectancy (e50) increase</td>
<td>0.165</td>
<td>0.079</td>
<td>114</td>
</tr>
<tr>
<td>Log sugars per capita</td>
<td>0.666</td>
<td>&lt;0.001</td>
<td>109</td>
</tr>
<tr>
<td>Log GDP per capita</td>
<td>0.720</td>
<td>0.720</td>
<td>116</td>
</tr>
<tr>
<td>Log BMI≥30 prevalence</td>
<td>0.636</td>
<td>&lt;0.001</td>
<td>109</td>
</tr>
<tr>
<td>Log urbanization</td>
<td>0.507</td>
<td>&lt;0.001</td>
<td>118</td>
</tr>
</tbody>
</table>

*Keeping GDP, BMI, urbanization and sugars intake constant. BMI, body mass index; GDP, gross domestic product; $I_{bs}$, Biological State Index.
Our study showed that the relationship between life expectancy and T1D prevalence rate was exponential (figure 2, $R^2=0.5266$). The Australian Institute of Health and Welfare (AIHW) also indicated the exponential relationship between T1D prevalence rate between age increase of Australian population through the Australian National Diabetes Register.50 Additionally, Neville et al31 reported that the increased longevity of diabetic patients contributed to the increasing prevalence of diabetes in Japanese population. The life expectancy gap between patients with TID and non-diabetic people has reduced significantly52 due to developments in sanitation, medical interventions and public health measures. Therefore, the underlying reason for the exponential relationship in our study may be because the number of individual patients with TID have increased in the human population.3 The American Diabetes Association has also stated that the majority of individuals with TID are adults even though TID has been more frequent and a relatively straightforward diagnosis in children.53

The correlations of T1D prevalence rates to both $I_{bs}$ and life expectancy were not only observed worldwide, but also in different country groupings sharing specific characteristics such as geographic locations (table 2), culture backgrounds (table 3) and affiliations to international functional organizations (table 3). Results’ highlights indicated that the correlations of $I_{bs}$ and life expectancy to T1D were significant or very strong in European country groupings (WHO-Europe in table 2 and EEA in table 3), but very weak in Asian country grouping (WHO-SEA in table 1 and ACD in table 3). This may be attributable to high-genetic predispositions13–15 40 in Europe, but low-genetic predisposition in Asia.46 53

Although we found that the correlations of $I_{bs}$ and life expectancy to T1D prevalence rate existed globally and in different country groupings categorized with a variety of criteria, there are several limitations, including the intrinsic limitations (conceptualized as ecological fallacy) to this study. First, the data analyzed were calculated for per capita in each country, so we could only demonstrate the relationships between TID prevalence rate and $I_{bs}$ and life expectancy at country/population level, which does not necessarily correspond to the same relationships holding true at the individual level. We also need to point out that it would be difficult to test the relationships at the individual level due to very rare TID occurrence rate.

Second, the slow changes in the genetic code of the human populations may not fully explain the increasing T1D prevalence. Non-genetic (environmental) factors partially determine whether, and how risk-associated genotypes may lead to overt T1D disease. Unfortunately, our $I_{bs}$ does not indicate if fitness change at population level is due to evolution of individuals or change/s. It may also be that altered lower natural immunity to infections following decades of using antibiotics may influence increased rates of autoimmune diseases including T1D.

Third, the data compiled and/or collected by the major international agencies (IDF, WHO, FAO and the World Bank) are fairly crude, and may contain some random errors.

Finally, current evidence of the increasing frequency of many heritable genetic disorders, including TID does not appear to be available. To the best of our knowledge, the TID prevalence rate for all age groups at country level published by IDF may be the only version to single out T1D prevalence worldwide after consulting the major diabetes research or data collecting institutions. This may be because clinically, adult TID is difficult to discriminate from certain forms of type 2 diabetes and from latent autoimmune diabetes in adults (LADA).53 Therefore, we do not know how much this set of data was confounded by other forms of diabetes.

The current prevailing paradigm on the increasing prevalence of T1D is that environmental pressures are now able to trigger genotypes.8 9 12–17 Currently, medical gene intervention in modern medicine at this stage cannot remove TID genes, and eugenics (improvement in the genetic stock) can offer no direction due to ethics issue. Therefore, study of TID epidemiology based on prevalence/incidence T1D data of all age groups has become imperative as it may offer optimal solution to address or at least slow down TID genetic load increases in different populations.

CONCLUSIONS

Our study suggested that reduced natural selection ($I_{bs}$) may be the major contributor to the increasing prevalence of T1D worldwide with special regard to European countries. It seems that TID epidemiology study based on all age groups may be the practical solution to identify the causes of increasing T1D prevalence and to address, or at least slow down, TID genetic load increases in different populations as modern medicine cannot operate effectively at the gene level yet.

Acknowledgements The authors express appreciation to Dr Arthur Saniotis for his editorial assistance.

Contributors MH conceived the idea for this study, W-PY extracted the data and W-PY took responsibility of data integrity. MH and W-PY interpreted the data. W-PY wrote the manuscript. W-PY and MH edited and approved the manuscript for submission to the journal.

Competing interests Financial support to MH was provided by the Wood Jones Bequest to the University of Adelaide.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Not relevant as all data for this research were freely downloaded from the United Nations agents’ websites. However, the corresponding authors could supply the country-specific data used in this research on request.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/
REFERENCES

30. The European Free Trade Association. What is the European Economic Area? http://www.efta.int/eea