Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to:

Historical HbA1c Values May Explain the Type 2 Diabetes Legacy Effect: UKPDS 88

Marcus Lind MD, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden; Department of Medicine, NU-Hospital Group, Uddevalla, Sweden

Henrik Imberg, Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, Gothenburg, Sweden

Ruth L. Coleman, Diabetes Trials Unit, Radcliffe Department of Medicine, University of Oxford, Oxford, U.K

Olle Nerman, Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, Gothenburg, Sweden

Rury R. Holman, Diabetes Trials Unit, Radcliffe Department of Medicine, University of Oxford, Oxford, U.K
Contents

Additional statistical analysis details........................................................................................................... 3

Tables ...................................................................................................................................................... 6

Table S1. Estimated parameters with 95% confidence intervals for the variables included in the final model (Equation 1). ...................................................................................................... 6

Table S2. Estimated relative risks of all-cause mortality and myocardial infarction between 0–10, 10–15 and 10–20 years after diagnosis assuming 0.5 or 2 percentage units (5.5 or 22 mmol/mol) lower HbA1c from diagnosis, and when the same HbA1c lowering was imposed from 5 and from 10 years after diagnosis. .................................................................................. 7

Figures ..................................................................................................................................................... 8

Figure S1. Estimated weight function $g(t)$ (Equation 3) of the influence weighted HbA1c variable (Equation 2) when analysing the time dependent effects of HbA1c on all-cause mortality (left) and myocardial infarction (right). The corresponding estimates of the parameters $b1$, $b2$ and $b3$ are provided in Table S1. .................................................................. 8

Figure S2. Time-dependent hazard ratios for all cause-mortality (left) and myocardial infarction (right) from 0 to 20 years after diagnosis of type 2 diabetes assuming a 0.5-percentage unit lower HbA1c from diagnosis (green dotted lines), and when the same HbA1c lowering was imposed from 5 years (blue dashed lines) and from 10 years (red solid lines) after diagnosis. The shaded regions represent 95% confidence limits. Hazard ratios were calculated according to Equation 6. ........................................................................................... 9

Figure S3. Time-dependent hazard ratios for all cause-mortality (left) and myocardial infarction (right) from 0 to 20 years after diagnosis of type 2 diabetes assuming a 2-percentage unit lower HbA1c from diagnosis (green dotted lines), and when the same HbA1c lowering was imposed from 5 years (blue dashed lines) and from 10 years (red solid lines) after diagnosis. The shaded regions represent 95% confidence limits. Hazard ratios were calculated according to Equation 6. ........................................................................................................................... 10

Figure S4. Cumulative number of events for observed and model predicted all-cause mortality (left) and myocardial infarction (right) during follow-up. ................................................................. 11
Additional statistical analysis details

Statistical analyses were performed using Poisson regression with current (updated) age, current (updated) diabetes duration, sex (male = 1, female = 2) and an influence weighted HbA1c variable (%) as explanatory variables. The total follow-up period of each patient was subdivided into small intervals of 0.2 years where a Poisson model with constant hazard was assumed (or, equivalently, an exponential distribution for the survival time). The subdivision into small intervals makes the assumption of Poisson distribution well fulfilled, allows for time dependent covariates and enables flexible modelling of the impact of historical HbA1c values at various points in time.

Model formulation

Using the Poisson model with piecewise constant hazard, the contribution to the likelihood function per interval of an individual is \((\lambda)^k \exp(-\lambda l_k)\), where \(k = 0\) or \(1\) depending on whether an event (only the individual’s first one) had occurred in the interval. The quantity \(l\) is the length of the contribution period in the interval (at most 0.2 years and shorter if there was an event or censoring within the interval), and \(\lambda\) is the hazard. In each time interval, the hazard \(\lambda(t)\) at time \(t\) was modelled as

\[
\log(\lambda(t)) = \beta_1 + \beta_2 \times \text{Age}(\tau) + \beta_3 \times \text{Diabetes duration}(\tau) + \beta_4 \times \text{Sex} + \beta_5 \times w\text{HbA1c}(\tau), \quad (Eq. 1)
\]

where \(\text{Age}(\tau)\), \(\text{Diabetes duration}(\tau)\) and \(w\text{HbA1c}(\tau)\) are the age, diabetes duration and influence weighted HbA1c at time \(\tau = 0.2 \times \lfloor 5t \rfloor\), i.e. evaluated at the left endpoint of the current time interval. The influence weighted HbA1c variable was defined as an integral of historical HbA1c values

\[
\int_0^t x(s)g(t-s)ds, \quad (Eq. 2)
\]

where \(x(s)\) is the HbA1c value at time point \(s\) (years since diagnosis) using linear interpolation between observed HbA1c values, and \(g(t)\) is a weight function. The weight function \(g(t)\) was defined as a piecewise exponential function with one knot:

\[
g(t) = \begin{cases} 
\exp(b_1 t) & \text{if } t \leq b_2 \\
\exp(b_1 b_2 + b_3 (t - b_2)) & \text{if } t > b_2
\end{cases}, \quad (Eq. 3)
\]

where \(b_1, b_2\) and \(b_3\) are parameters to be estimated. These parameters may be interpreted as follows: \(b_1\) describes an initial increase or decrease in the relative risk contribution over time from an HbA1c value, and \(b_2\) describes the increase or decrease of the relative risk contribution after the breakpoint \(b_2\). The shape of the function \(g(t)\) for the outcomes considered in this study and with the parameters \(b_1, b_2\) and \(b_3\) estimated from data is presented in Figure S1.

Time-dependent HbA1c hazard ratios and relative risks

Consider two continuous HbA1c curves \(\{x_0(s), s \in [0, T]\}\) and \(\{x_1(s), s \in [0, T]\}\) on a time interval from 0 to \(T\) years after diagnosis, where \(x_0(s)\) and \(x_1(s)\) are the HbA1c values at the time point \(s\) \(\in [0, T]\). We describe below how hazard ratios and relative risks of the HbA1c profile \(x_1\) vs \(x_0\) may be calculated from the Poisson model with the hazard function defined by Equation 1–3.

Hazard ratio between two HbA1c profiles

According to Equation 1 and 2, the hazard ratio of the HbA1c profile \(x_1\) vs \(x_0\) at time \(t\) is given by
If \( x_1(s) = x_0(s) + z \) for all \( s \in [0,T] \) and some constant \( z \), i.e. for a constant shift in HbA1c, Equation 4 simplifies to

\[
HR(t) = e^{\beta z \int_0^t g(s) ds}. \quad (Eq. 5)
\]

In particular, the hazard ratio for a constant shift in HbA1c is independent of the reference HbA1c profile \( x_0 \). If, on the other hand, \( x_1 \) is given by

\[
x_1(s) = x_0(s) + z \times 1_{(s > t_0)} = \begin{cases} 
x_0(s), & s \leq t_0 \\
x_0(s) + z, & s > t_0
\end{cases}
\]

i.e. the shift is imposed first at time \( t_0 \), the hazard ratio function becomes

\[
HR(t) = \begin{cases} 
1, & t \leq t_0 \\
e^{\beta z \int_{t_0}^t g(s) ds}, & t > t_0
\end{cases}. \quad (Eq. 6)
\]

The cumulative weight ascribed to HbA1c values the first \( s \) years after diagnosis to the effect of HbA1c on the hazard \( t \) years after diagnosis is given by

\[
\int_{t-s}^t g(u) du / \int_0^t g(u) du. \quad (Eq. 7)
\]

**Relative risk between two HbA1c profiles**

The survival function \( S(t) = Prob(No \text{ event before time } t) \) can be calculated from the hazard function \( \lambda(t) \) according to the formula

\[
S(t) = e^{-\Lambda(t)} \approx 1 - \Lambda(t), \quad (Eq. 8)
\]

where

\[
\Lambda(t) = \int_0^t \lambda(s) ds
\]

is the cumulative hazard function. The approximation in Equation 8 follows from a Taylor expansion of the exponential function and is appropriate for events with low probabilities. The risk of an event in a time interval \([s, t]\) from \( s \) to \( t \) years after diagnosis is thus given by

\[
Prob(\text{Event time in interval } [s, t]) = 1 - S(t) - (1 - S(s)) \approx \Lambda(t) - \Lambda(s)
\]

\[
= \int_s^t \lambda(u) du. \quad (Eq. 9)
\]

The absolute risk depends on all the covariates in the model and on the time interval of interest.

Considering two different patient and HbA1c profiles, the relative risk of an event in the time interval \([s, t]\) can be calculated as the ratio of corresponding absolute risks obtained from Equation 9, i.e.

\[
\int_s^t \lambda_1(u) du / \int_s^t \lambda_0(u) du, \quad (Eq. 10)
\]

where \( \lambda_1(u) \) and \( \lambda_0(u) \) are the corresponding hazard functions. Keeping the other covariates fixed, the relative risk due to differences in two HbA1c curves \( x_1 \) and \( x_0 \) becomes a function of the time since diagnosis, reference HbA1c profile \( x_0 \) and hazard ratio function \( HR(t) \) (Equation 4) between the two.
HbA1c profiles. When evaluating the relative risks associated with various HbA1c profiles we found the relative risk to be essentially independent of the reference HbA1c profile $x_0$, motivating the approximation

$$\frac{1}{t-s} \int_s^t HR(u) du \quad (Eq. 11)$$

of the relative risk of an event in a time interval $[s, t]$.

**Parameter estimation and hypothesis testing**

Estimation was performed using maximum likelihood, where the parameters $b_1, b_2$ and $b_3$ were estimated simultaneously with the regression coefficients $\beta_1, ... , \beta_5$; a possibility offered by the use of Poisson regression instead of e.g. Cox regression. The significance of individual regression coefficients was assessed by likelihood ratio tests, and corresponding confidence intervals were computed by test inversion. Estimates and confidence intervals for the hazard ratio of the influence weighted HbA1c variable at various follow-up times and for the relative risk associated with early and late HbA1c reductions were computed from the corresponding regression coefficient ($wHbA1c$, Table S1), fixing the parameters of the HbA1c weight function $g(t)$ at their estimated values.
### Tables

**Table S1.** Estimated parameters with 95% confidence intervals for the variables included in the final model (Equation 1).

| Regression coefficients | Parameter estimate (95% CI) | All-cause mortality | Myocardial infarction |
|-------------------------|-------------------------------|---------------------|-----------------------|
| **Intercept**           |                               | -9.7094 (-9.510 – -9.935) | -7.4416 (-7.692 – -7.223) |
| **Current diabetes duration (years)** |                               | -0.1318 (-0.152 – -0.114) | -0.1275 (-0.1537 – -0.1052) |
| **Current age (years)** |                               | 0.09953 (0.0962 – 0.1024) | 0.0641 (0.0602 – 0.0675) |
| **Sex (male = 1, female = 2)** |                               | -0.53154 (-0.6846 – -0.3785) | -0.7239 (-0.8968 – -0.5510) |
| **wHbA1c**              |                               | 0.0032475 (0.0027 – 0.0037) | 0.0832 (0.0682 – 0.0960) |
| **Weight function g(t)** |                               |                      |                       |
| **b_1**                 |                               | 0.704737 (0.638 – 0.755) | -5.2662 (-6.310 – -4.507) |
| **b_2**                 |                               | **2.7417 (2.47 – 2.95)** | **0.1967 (0.1673 – 0.2377)** |
| **b_3**                 |                               | -0.03853 (-0.080 – -0.012) | 0.1967 (0.1673 – 0.2377) |

wHbA1c is the influence weighted HbA1c variable (Equation 2) using the influence (weight) function $g(t)$ (Equation 3).

- $b_1$ describes the initial increasing/decreasing phase of the function $g(t)$.
- $b_2$ is the breakpoint of the piecewise exponential function $g(t)$.
- $b_3$ describes the increase/decrease of the function $g(t)$ after the breakpoint $b_2$.

CI, confidence interval; HbA1c, Hemoglobin A1c.
Table S2. Estimated relative risks of all-cause mortality and myocardial infarction between 0–10, 10–15 and 10–20 years after diagnosis assuming 0.5 or 2 percentage units (5.5 or 22 mmol/mol) lower HbA1c from diagnosis, and when the same HbA1c lowering was imposed from 5 and from 10 years after diagnosis.

| Years after diagnosis | Relative risk (95% CI) \(per 0.5\) percentage units lower HbA1c | | Relative risk (95% CI) \(per 2\) percentage units lower HbA1c |
|-----------------------|---------------------------------------------------------|---------------------------------------------------------|
|                       | HbA1c lowered at diagnosis | HbA1c lowered 5 years after diagnosis | HbA1c lowered 10 years after diagnosis | HbA1c lowered at diagnosis | HbA1c lowered 5 years after diagnosis | HbA1c lowered 10 years after diagnosis |
| **All-cause mortality** |                                         |                                         |                                         |                                         |                                         |                                         |
| 0–10                  | 0.963 (0.958 - 0.969) | 0.993 (0.992 - 0.995) | 1.00 | 0.864 (0.847 - 0.884) | 0.973 (0.970 - 0.977) | 1.00 |
| 10–15                 | 0.902 (0.889 - 0.917) | 0.941 (0.933 - 0.950) | 0.987 (0.985 - 0.989) | 0.660 (0.623 - 0.705) | 0.782 (0.756 - 0.814) | 0.946 (0.939 - 0.955) |
| 10–20                 | 0.886 (0.871 - 0.903) | 0.921 (0.910 - 0.933) | 0.963 (0.958 - 0.969) | 0.616 (0.576 - 0.665) | 0.721 (0.689 - 0.759) | 0.864 (0.847 - 0.884) |
| **Myocardial infarction** |                                         |                                         |                                         |                                         |                                         |                                         |
| 0–10                  | 0.945 (0.936 - 0.954) | 0.984 (0.981 - 0.987) | 1.00 | 0.799 (0.772 - 0.831) | 0.937 (0.929 - 0.948) | 1.00 |
| 10–15                 | 0.896 (0.881 - 0.914) | 0.923 (0.911 - 0.936) | 0.967 (0.962 - 0.973) | 0.644 (0.602 - 0.697) | 0.723 (0.688 - 0.767) | 0.875 (0.857 - 0.896) |
| 10–20                 | 0.888 (0.872 - 0.907) | 0.909 (0.896 - 0.925) | 0.945 (0.936 - 0.954) | 0.622 (0.578 - 0.677) | 0.683 (0.645 - 0.732) | 0.799 (0.772 - 0.831) |

The relative risk of an event in a time interval 0–10, 10–15 or 10–20 years after diagnosis was calculated according to Equation 11.

CI, confidence interval; HbA1c, Hemoglobin A1c.
Figures

**Figure S1.** Estimated weight function $g(t)$ (Equation 3) of the influence weighted HbA1c variable (Equation 2) when analysing the time dependent effects of HbA1c on all-cause mortality (left) and myocardial infarction (right). The corresponding estimates of the parameters $b_1$, $b_2$ and $b_3$ are provided in Table S1.
**Figure S2.** Time-dependent hazard ratios for all cause-mortality (left) and myocardial infarction (right) from 0 to 20 years after diagnosis of type 2 diabetes assuming a 0.5-percentage unit lower HbA1c from diagnosis (green dotted lines), and when the same HbA1c lowering was imposed from 5 years (blue dashed lines) and from 10 years (red solid lines) after diagnosis. The shaded regions represent 95% confidence limits. Hazard ratios were calculated according to Equation 6.
Figure S3. Time-dependent hazard ratios for all cause-mortality (left) and myocardial infarction (right) from 0 to 20 years after diagnosis of type 2 diabetes assuming a 2-percentage unit lower HbA\(_1c\) from diagnosis (green dotted lines), and when the same HbA\(_1c\) lowering was imposed from 5 years (blue dashed lines) and from 10 years (red solid lines) after diagnosis. The shaded regions represent 95% confidence limits. Hazard ratios were calculated according to Equation 6.
**Figure S4.** Cumulative number of events for observed and model predicted all-cause mortality (left) and myocardial infarction (right) during follow-up.