

OPEN

Jointly modelling longitudinally measured urinary human chorionic gonadotrophin and early pregnancy outcomes

N. B. Ashra^{1*}, L. Marriott², S. Johnson², K. R. Abrams¹ & M. J. Crowther¹

Human chorionic gonadotrophin (hCG) is largely used to confirm pregnancy. Yet evidence shows that longitudinal hCG profiles are distinguishable between healthy and failing pregnancies. We retrospectively fitted a joint longitudinal-survival model to data from 127 (85 healthy and 42 failing pregnancies) US women, aged 18–45, who were attempting to conceive, to quantify the association between longitudinally measured urinary hCG and early miscarriage. Using subject-specific predictions, obtained uniquely from the joint model, we investigated the plausibility of adaptively monitoring early pregnancy outcomes based on updating hCG measurements. Volunteers collected daily early morning urine samples for their menstrual cycle and up to 28 days post day of missed period. The longitudinal submodel for log hCG included a random intercept and slope and fixed linear and quadratic time terms. The survival submodel included maternal age and cycle length covariates. Unit increases in log hCG corresponded to a 63.9% (HR 0.36, 95% CI 0.16, 0.47) decrease in the risk of miscarriage, confirming a strong association between hCG and miscarriage. Outputted conditional survival probabilities gave individualised risk estimates for the early pregnancy outcomes in the short term. However, longer term monitoring would require a larger sample size and prospectively followed up data, focusing on emerging extensions to the joint model, which allow assessment of the specificity and sensitivity.

Early miscarriage, defined in the UK as loss before week 13, is a frequent complication of pregnancy¹. It affects 12% to 24% of clinically confirmed pregnancies, not counting those losses which occur prior to the date of the missed period - so-called biochemical pregnancies². Women who suffer from a miscarriage are more likely to report symptoms associated with depression, with affected women ranging from 20% to a high of 55%³. Though the majority of losses are self-resolving, those that are not may require diagnostic tests, hospital treatment, surgical intervention and follow-up care². This provides an incentive to identify potential early losses as early as possible by exploring more patient-centred monitoring strategies.

The recently published priorities for research within miscarriage ranked highest the identification of effective interventions to prevent miscarriage⁴. This encompasses the plausibility of using biomarkers to track pregnancy progression through viability or miscarriage. Several potential biomarkers have been identified to predict miscarriage, with human chorionic gonadotrophin (hCG) a strong contender⁵. The hormone tends to rise rapidly and reliably in early pregnancy, doubling every 1.5 days in the first 5 weeks post conception and then every 3.5 days from week 7, before plateauing around week 10^{5,6}. Its use is more prevalent in tracking early pregnancy progress in an *in vitro* fertilisation (IVF) population and for identifying ectopic pregnancies⁷. However, evidence suggests that longitudinal profiles of hCG can be utilised to distinguish between viable and failing pregnancies, with similar patterns of hCG noted across maternal serum and urine⁸.

The repeated collection of a continuous biomarker, such as hCG, over time gives rise to intermittently observed longitudinal data which are subject to measurement error^{9,10}. Conventionally, this data is analysed using linear mixed effects models, with time-to-event outcomes analysed using survival models^{11,12}. However, when interest lies in quantifying the association between the repeatedly measured biomarker and time-to-event outcome, separate analyses ignore the dependency between the longitudinal and time-to-event processes¹³.

¹Department of Health Sciences, University of Leicester, Leicester, UK. ²Clinical Research Department, SPD Development Company Ltd., Bedford, UK. *email: nbba1@le.ac.uk

Variables	Healthy (n = 85)	Miscarried (n = 44)	Overall (n = 129)
Age, years	29.95 (4.15)	32.34 (4.60)	30.77 (4.44)
Ethnicity, n (%)			
White	75 (88.24)	34 (77.27)	109 (84.50)
Black	3 (3.53)	7 (15.91)	10 (7.75)
Asian	4 (4.71)	2 (4.55)	6 (4.65)
Mixed	3 (3.53)	1 (2.27)	4 (3.10)
Education, n (%)			
High School	4 (4.71)	2 (4.55)	6 (4.65)
Graduate	69 (81.18)	28 (63.64)	97 (75.19)
Postgraduate	12 (14.12)	14 (31.82)	26 (20.16)
Occupation, n (%)			
Homemaker	12 (14.12)	3 (6.82)	15 (11.63)
Student	1 (1.18)	1 (2.27)	2 (1.55)
Skilled labourer	2 (2.35)	2 (4.55)	4 (3.10)
Office admin	8 (9.41)	5 (11.36)	13 (10.08)
Professional	60 (70.59)	31 (70.45)	91 (70.54)
Other	2 (2.35)	2 (4.55)	4 (3.10)
Cycle length	29.94 (2.95)	28.66 (3.21)	29.50 (3.09)
Previous pregnancies	1.00 (1.05)	1.11 (1.15)	1.04 (1.08)
Previous live births	0.62 (0.76)	0.70 (0.88)	0.65 (0.80)
Time to conceive, months	4.36 (5.83)	4.55 (5.98)	4.43 (5.86)
Previous miscarriage, n (%)	11 (12.94)	4 (9.76)	15 (11.90)

Table 1. Baseline demographics by pregnancy viability group. All values are mean(SD) unless otherwise stated.

Acknowledging an association between a longitudinal biomarker and survival outcome implies that very high or low values of the biomarker are indicative of adverse outcomes¹⁴. Fitting a simple survival model to the event, including all of the longitudinal biomarker information, tells us how a change in biomarker value affects survival over follow-up time. However, the variation in biomarker observations between individuals is not incorporated into the model, so inferences for individuals cannot be drawn. Secondly, the implicit changes in biomarker values between each physically observed measurement are ignored, resulting in a failure to build a complete biomarker profile. The linear mixed effects model can build this biomarker trajectory and inbuilt random effects allow estimation of personalised risk as an output of the model. Recognizing the advantages of both types of model, combining both the linear mixed effects and survival models through a shared dependence structure via the joint longitudinal-survival model is essential. This allows the association to be appropriately modelled, whilst taking into account the intermittent nature of observations and measurement error. The model, through estimation of individual trajectories, can aid monitoring and potentially prediction of outcomes.

The aim of this paper is to retrospectively apply the classical joint model framework to data of pregnant women, who were followed up from before conception, to quantify the association between longitudinal urinary hCG observations and early miscarriage. The paper will also consider whether estimation of conditional survival probabilities from the joint model could provide the basis for dynamic monitoring of patients in the very early stages of pregnancy prior to other symptoms manifesting.

Results

A total of 44 (17.6%) women suffered miscarriages. The dataset used for analysis consists of 85 randomly selected viable pregnancies and 44 miscarried pregnancies. A summary of demographic variables is given in Table 1. Overall, the two groups were comparable. Women who experienced healthy pregnancies were slightly younger (mean \pm SD: 29.95 \pm 4.15) than those who miscarried (mean \pm SD: 32.34 \pm 4.60). The majority of women in either group were from a White European background (88.24% and 77.27% respectively). A slightly higher proportion of women who had viable pregnancies had previously experienced a miscarriage, compared to women who miscarried (12.94%, and 9.76%). Of the women who miscarried, 18 (14.2%) experienced biochemical pregnancies and 24 (57.1%) women suffered early miscarriages. Two women who miscarried did not contribute hCG measurements and were not included in the joint modelling analysis.

The remaining 127 women all contributed repeated hCG measurements. For women who miscarried the average number of hCG observations was 17.5 (SD: 8.6) and for women who experienced viable pregnancies the average number of measurements was higher at 23.6 (SD 3.9).

Profiles of log hCG measurements for viable and failing pregnancies are presented in Fig. 1. The general trajectory shows an initial rise after conception, which continues through the first three weeks of the healthy pregnancies before slowing in rise. There was greater variation in profiles for women who miscarried, who also

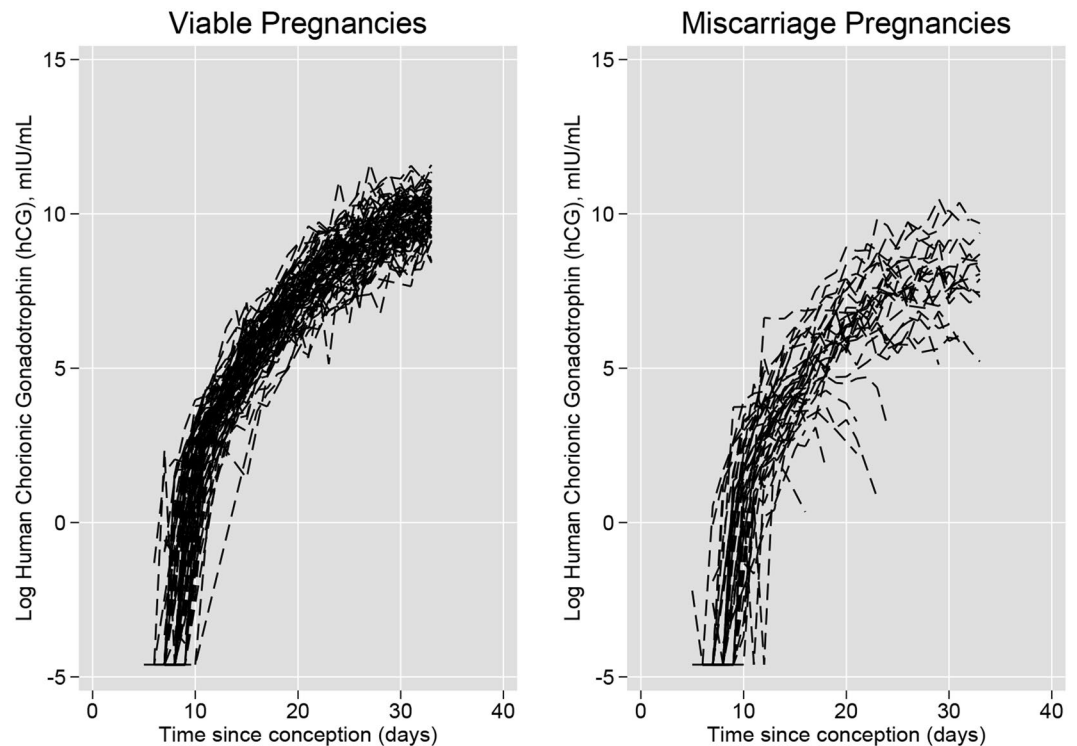


Figure 1. Log human chorionic gonadotrophin trajectories for viable pregnancies and miscarriage pregnancies.

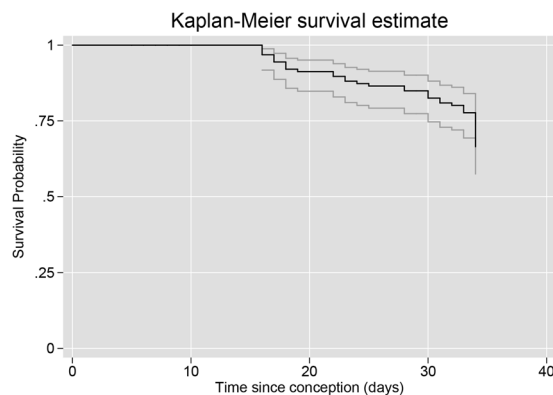


Figure 2. Kaplan-Meier survival probabilities for time-to-miscarriage.

presented with an initial rise after conception. However, some women experienced a sharp drop in hCG, whilst others experienced a more gradual rise in hCG in comparison with women who had healthy pregnancies.

Overall Kaplan-Meier survival estimates for time to miscarriage are shown in Fig. 2. An approximate 15 day lag is evident before an event is seen, due to the use of time since conception as a timeline.

Modelling longitudinal profile. Inclusion of a quadratic time variable was necessary to appropriately capture the shape of the log hCG profile. Results from an initial fitted linear mixed effects model, including a grouping variable for pregnancy outcome, confirmed that mean log hCG was -1.66 mIU/mL (95% CI -2.14 , -1.18) lower in the biochemical pregnancy group and -1.13 mIU/mL (95% CI -1.48 , -0.78) lower in the early miscarriage group, when compared with the healthy pregnancies. Results are presented in Table 2.

Joint longitudinal-survival model. A joint longitudinal-survival model was fitted to the data. Estimates for the model with current value association structure are given in Table 3. A unit increase in absolute value of log hCG corresponded to a 66.1% (HR 0.339, 95% CI 0.257, 0.447) decrease in the risk of miscarriage at time t . A one-year increase in maternal age at conception resulted in a 7.6% (HR 1.076, 95% CI 0.998, 1.159) increase in the risk of miscarriage. A one-day increase in cycle length was associated with a 15.6% (HR 0.844, 95% CI 0.739, 0.965) decrease in the risk of miscarriage.

Longitudinal model	Mean change in log hCG Mtu/ml	95% Confidence Interval
Time since conception, days	1.431	1.396, 1.466
Quadratic time since conception, days	-0.025	-0.026, -0.024
Group		
Healthy	—	—
Biochemical loss	-1.656	-2.135, -1.176
Early loss	-1.132	-1.484, -0.781

Table 2. Model estimates from a linear mixed effects model.

Survival submodel	Hazard Ratio	95% Confidence Interval
Age, years	1.076	0.998, 1.159
Usual cycle length, days	0.844	0.739, 0.965
Expected current value of log hCG	0.339	0.257, 0.447
Longitudinal submodel	Mean	95% Confidence Interval
Time since conception, days	1.431	1.396, 1.466
Quadratic time since conception, days	-0.025	-0.026, -0.024

Table 3. Model estimates from a joint longitudinal-survival model with current value association structure.

Model	Standard error	Hazard Ratio for log hCG	95% Confidence Interval
Time-varying covariate	0.036	0.439	0.373, 0.516
Two-stage model	0.040	0.440	0.368, 0.527
Joint model	0.142	0.339	0.257, 0.447

Table 4. Survival estimates from a standard survival model with time-varying covariate, two-stage model and joint model.

A comparison of log hCG association parameters for various models are presented in Table 4. The association between log hCG and time to miscarriage was attenuated when fitting a survival model with time-varying covariate (HR: 0.439, 95% CI: 0.373, 0.516) and the two-stage model (HR: 0.440 95% CI: 0.368, 0.527). Furthermore, standard errors for both the standard survival model and two-stage model were 0.036 and 0.040 respectively compared to a larger 0.142 for the joint model.

Conditional survival predictions. Conditional survival probabilities were obtained from the joint model, which included the current value association structure (see Table 2). Probabilities estimated for the ten-day window after the last observed hCG measurement are shown in Fig. 3 and for a two-day window in Fig. 4. Participants A and B experienced biochemical and early losses respectively, whilst participant C experienced a healthy pregnancy. For both participants A and B a similar number of measurements were observed over comparable time periods, with similar average cycle lengths (28 and 30 days respectively) and ages (42 and 38 years respectively). Based on observed hCG measurements as well as age and cycle information, both were predicted to experience miscarriages.

For participant C estimates confirmed an 80% survival probability for the pregnancy two days post last observed hCG measurement. Depending on the cut-off used for low risk this may not be considered a high enough survival probability for a healthy pregnancy. As follow-up did not continue it was not possible to update probabilities to look at longer-term outcomes.

Discussion

Principal findings. This analysis builds on the two-stage model approach implemented by Marriott *et al.*¹⁵. By utilising the more advanced joint longitudinal-survival framework, the association between longitudinally measured urinary hCG and time to miscarriage is modelled, accounting for both measurement error and the intermittent nature of observations. This improves upon the two-stage model, which assumed that measurements remained constant between observation times.

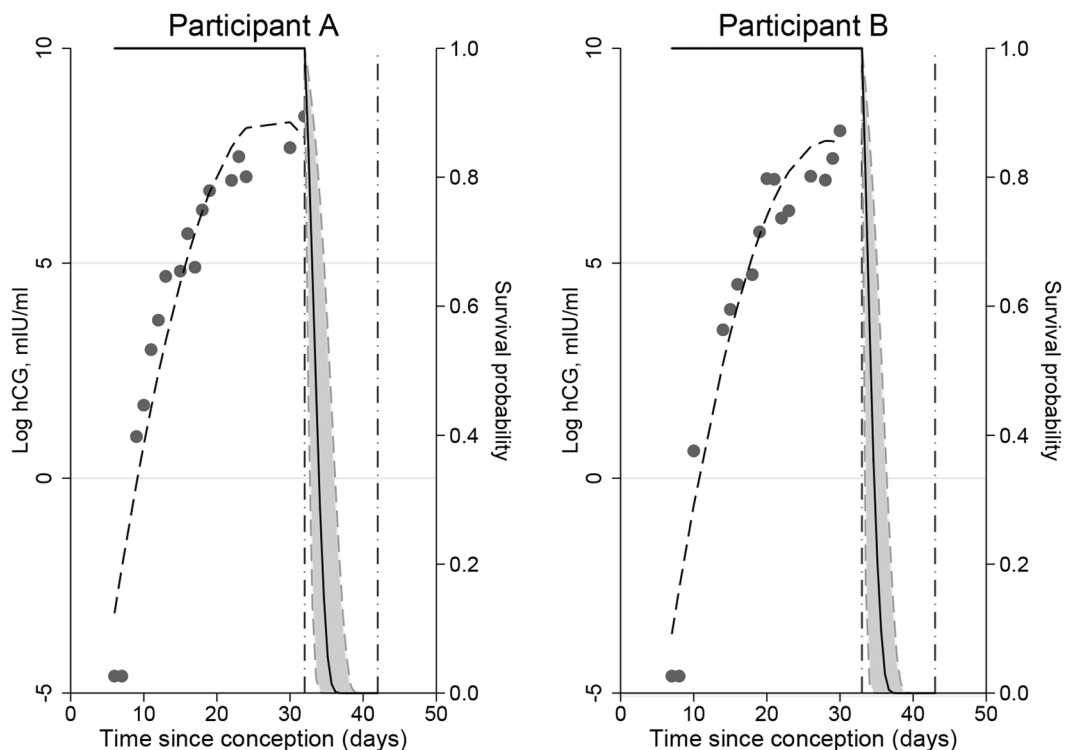


Figure 3. Conditional survival probability curves for participants A and B who experienced biochemical and early losses, respectively.

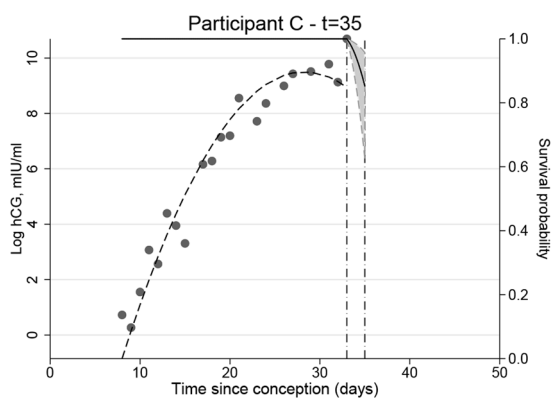


Figure 4. Conditional survival probability curve for participant C who experienced a healthy pregnancy.

With the emphasis now on personalised care, it is becoming standard practice to use the joint model in favour of singular or two-stage analyses to model the association between longitudinal and failure processes, to both maximise efficiency and minimise the potential for bias¹⁶. The mainstream use of joint models coincides with improvements in software making these complicated models increasingly easier to fit, with packages available in both R (JM, JoineRML) and Stata (*stjm*, *merlin*)^{13,17–19}. This makes the estimation of conditional survival probabilities from such models more accessible.

This paper investigates whether urinary hCG could be used to monitor pregnancy viability prospectively in early pregnancy from first detection of hCG. Tracking at this early stage presents an adjunct to diagnosis by ultrasound later on in the pregnancy. This analysis echoes research suggesting declines in hCG can be noted even prior to other symptoms presenting²⁰. There is also potential for this monitoring to occur prior to conception, with a recent study finding that a lag between the luteal phase and hCG production can be indicative of a biochemical pregnancy, possibly due to early or delayed implantation²¹.

Tracking of hCG by pregnant women is practicable, as demonstrated by Foo *et al.* who employed a fertility monitor that also provide semi-quantitative analysis of hCG levels on pregnancy tests that were used daily in women who conceived²¹. Retrospective analysis of the semi-quantitative data indicated that non-viable pregnancies had different hCG profiles to viable pregnancies. Serial tracking could have the potential to cause stress,

although women using tests to track ovulation for fertility purposes do not appear to have higher stress levels than those not employing tests^{22,23}. Nevertheless, it is likely that tracking would initially be of benefit in high risk pregnancies, where anxiety levels are already high and there would be a willingness and reason to track. Further research would be required to understand the psychological impact of tracking.

Monitoring from first detection has the potential to be useful in cases of recurrent miscarriage, particularly as research into treatment gains traction. A recently published feasibility study assessing the effectiveness of the diabetes drug sitagliptin as a treatment for recurrent miscarriage, presented promising findings²⁴. This trial builds on previous research, which found that in some cases of recurrent miscarriage it is the deterioration of stem-like cells in the uterus which contribute to pregnancy loss. When adjusted for age and baseline colony forming unit (CFU) counts, the CFU count was higher (RR 1.52, 95% CI 1.32, 1.75) in the sitagliptin group compared to placebo, pointing to successful regeneration of cells. These findings could revolutionise treatment for unexplained recurrent miscarriage, particularly as the more established progesterone therapy has not been shown to significantly impact the rate of live births (PROMISE and PRISM trials)^{25,26}.

Not all miscarriage is likely to be predictable due to the diverse aetiology of the condition. Some causes can be directly related to reduced hCG levels, e.g. conditions that affect rate of embryonic development such as chromosomal abnormalities, or inadequate placentation. Other causes, for example, where infectious agents or trauma are involved, may have no forewarning.

The demographic factors that add to the model have plausibility. The association between chronological age and miscarriage is well documented, and short follicular phase has also been associated with miscarriage by other authors^{27,28}. Short cycle length may represent a surrogate marker for advanced reproductive age as the initial transition to peri-menopause can be characterised by a shortening of cycle length²⁹.

Strengths and limitations. The two-stage model did not allow the investigation of the nature of the association between miscarriage and hCG, something which is possible with the joint model. Although attempted it was not feasible to sensibly fit a joint model with a first derivative association structure, possibly due to the small sample size. This is something that requires further investigation in a larger dataset, particularly as there is evidence in the literature, which suggests that the overall profile of hCG is important as opposed to changes in absolute values of hCG. Certainly, both recent papers utilising Bayesian non-parametric models, and mixed effects penalized splines model approaches, focused on classification of each type of pregnancy based on complete longitudinal profiles^{30,31}.

A Weibull model was utilised to model the baseline hazard, however ideally more flexibility would be desirable. This could be achieved by using restricted cubic splines to model the baseline hazard. Model selection was carried out using forwards stepwise selection, which is known to introduce bias³². Alternative selection methods should be considered in future, specific to the joint modelling context. Selection based on the log likelihood contribution for the longitudinal part and conditional survival model have been proposed, but are currently only implemented in the SAS statistical software³³. The example dataset was relatively small, and so fitting a model as complex as the joint model was challenging. Results must therefore be interpreted with caution. As this was a retrospective analysis of data with limited follow-up measurements, it was not possible to update predictions as measurements were observed. Predictions were therefore inaccurate for wider time periods. With the small sample size, it was also not viable to split the dataset for development and validation of the model. Attempting to utilise such data for diagnostic or monitoring purposes also requires careful consideration of the potential for false positives. This study did not take into account the sensitivity and specificity of the fitted model, however this is an important component for planned future analyses in line with developments in joint model methodology^{34,35}.

When utilising the joint model framework, it is essential to think about adjustments that need to be made to the model to truly reflect the biological reality of the biomarker and disease processes. In this analysis considerations were made for the timeline on which miscarriage was modelled and how this affected the inclusion of fixed and random effects. Due to limitations of the software, the models included fixed and random intercepts, though no hCG would have been detectable at time zero. Date of conception would also be unknown in a natural pregnancy setting, making this analysis more suited to an IVF setting. This, however, could be adjusted for by using the last menstrual period (LMP) as a timeline in a natural pregnancy setting.

Employing two separate modelling techniques for longitudinal and survival data requires larger sample size requirements in a clinical trial setting. The increased efficiency of simultaneously modelling the two outcomes has the advantage of maintaining desired power at a lower sample size³⁶. This makes designing clinical trials around a joint model framework an attractive prospect.

Conclusions

The novel extension to this analysis concerns the subject-specific predictions. This study is an initial investigation into whether women at high risk of miscarriage could be adaptively monitored via their urinary hCG concentration. Though the effectiveness of possible treatments, particularly for recurrent miscarriage, remain uncertain; the joint model is well placed for dynamic monitoring. However long term follow-up observations are required, along with access to a larger dataset for a model to be developed and subsequently validated. Future analyses should also consider the sensitivity and specificity of the fitted predictive model, to minimise the likelihood of false diagnoses of miscarriage.

Methods

Description of dataset. Women attempting to conceive were asked to collect daily early morning urine samples for their entire menstrual cycle and up to 28 days after the day of their missed period if they became pregnant. Women recruited were aged between 18–45 years and were not excluded on the basis of existing fertility issues. Intra-individual variation in the concentration of first morning urine is much lower, than when

considering all urine voids. In addition, the exponential rise of hCG in early pregnancy from <1 mIU/ml to >150 000 mIU/ml renders fluctuations in urine concentration as having minimal effect on the trajectory of rise. Therefore, correction for urine concentration differences, e.g. using creatinine, was not required.

Urinary concentration of hCG was quantified using a validated quantitative immunoassay system (AutoDELFIA; PerkinElmer, Waltham, USA). The concentration of luteinising hormone (LH) in the urine were analysed by a panel of experts from a range of disciplines, including statisticians, endocrinologists and clinical scientists, to determine the day of LH surge, which occurs approximately 24 hours prior to ovulation. It was assumed conception occurred the day after the LH surge³⁷. Details of sample collection and storage have been described previously¹⁵. Additional maternal demographics, menstrual and pregnancy history data were recorded. The study was carried out in accordance with the ethical principles of the Declaration of Helsinki. Written, informed consent was obtained from all individual participants involved in the study.

Statistical analyses. The data utilised in this analysis have been analysed previously using a two-stage model approach¹⁵. The two-stage model utilises existing modelling techniques by first fitting a linear mixed effects model to the longitudinal data. Subject-specific predictions are then obtained from the mixed model and included as a time-varying covariate in a survival model. This method incorrectly assumes that the biomarker remains stagnant between measurements and gives too precise estimates with unrealistically small standard errors¹⁶. This analysis will now be extended using a joint model framework, which offers a number of advantages over the two-stage model approach.

Longitudinal modelling. The joint model is made up of two component models – the longitudinal linear mixed effects model and proportional hazards survival model³⁶. The longitudinal model for urinary log hCG forms a trajectory function, which estimates the unobserved values of log hCG for the i^{th} patient at time t to form complete profiles. The formulation of the fitted model is as follows,

$$\log hCG_i(t) = m_i(t) + e_{ij}$$

$$m_i(t) = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i}) \text{ time} + \beta_2 \text{ time}^2$$

$$b_i \sim MVN(0, \Omega_u) \quad e_{ij} \sim N(0, \sigma_e^2)$$

The model is made of fixed effects parameters including a fixed intercept (β_0) and linear and quadratic time since conception terms, with parameter estimates, β_1 and β_2 respectively. The random effects parameters allow each individual i to vary at baseline via a random intercept (b_{0i}) and over time through a random linear time since conception term, with parameter estimate b_{1i} . The possibility of measurement error, as with any continuous biomarker, is accounted for via the residual error term, e_{ij} , which is normally distributed. The random effects b_i are multivariate normally distributed. An unstructured correlation matrix was assumed.

Survival modelling. A proportional hazards survival submodel was assumed, conditional on $M_i(t) = \{m_i(s), 0 \leq s \leq t\}$, which denotes the history of the true unobserved longitudinal measurements up to time t and additional covariates v_i ¹⁴. The specific fitted model is given by,

$$h(t|M_i(t), v_i) = h_0(t) \exp[\gamma_1 \text{age} + \gamma_2 \text{usual cycle length} + \alpha m_i(t)]$$

The baseline hazard, $h_0(t)$, was assumed to follow a Weibull distribution. Maternal age and usual cycle length were included as covariates in the survival model, after a forwards model selection procedure was carried out at the 5% significance level. The inclusion of $m_i(t)$ in the survival submodel estimates the change in absolute log hCG values and is termed the current value parameterisation, with association parameter α . By including the longitudinal model within the survival submodel, we effectively link the expected value of log hCG to the miscarriage or censoring time, where typically an hCG response would not have been observed. Various association structures were explored, including the first derivative association structure, which models the rate of change of log hCG.

To allow for comparison a standard survival model with log hCG included as a time-varying covariate was fitted, as well as a two-stage model using subject specific predictions from the longitudinal model, as defined for the joint model, in a survival model.

Subject-specific survival probabilities dependent on maternal age and longitudinal log hCG measurements were obtained from the sample on which the joint model was fitted, using the Stata package `stjm`¹³. Conditional survival predictions can potentially be updated as measurements are observed, giving a real-time risk of miscarriage, or dynamic predictions. All models were fitted in Stata IC version 15.1.

Funding and ethical approval. This was a diagnostic accuracy study on a sample bank collected from a multicentre, prospective study, conducted by Radiant Research (USA) on behalf of the sponsor SPD Development Company Ltd. (UK). The study was approved by Quorum Review Committee on 30th November 2009; clinical trial number NCT01077583. This analysis was conducted by N.B.A as part of a doctoral training programme jointly funded by MRC IMPACT and SPD Development Company Ltd.

Data availability

The datasets analysed during the current study are not publicly available due to confidentiality restrictions in place between SPD Development Company Ltd. and the University of Leicester. Stata code written/used to perform the analysis are available from the corresponding author on reasonable request.

Received: 18 September 2019; Accepted: 26 February 2020;

Published online: 12 March 2020

References

- National Institute for Health and Care Excellence. Scenario: Managing suspected miscarriage. <https://cks.nice.org.uk/miscarriage#scenario> (2018).
- Jurkovic, D., Overton, C., & Bender-Atik, R. Diagnosis and management of first trimester miscarriage. *BMJ*. **346** (2013).
- Robinson, G. E. Pregnancy loss. *Best. Pract. Res. Clin. Obstet. Gynaecology*. **28**, 169–178 (2014).
- Prior, M. *et al.* Priorities for research in miscarriage: a priority setting partnership between people affected by miscarriage and professionals following the James Lind Alliance methodology. *BMJ Open*. **7**, e016571, <https://doi.org/10.1136/bmjopen-2017-016571> (2017).
- Barnhart, K. T. *et al.* Symptomatic Patients With an Early Viable Intrauterine Pregnancy: hCG Curves Redefined. *Obstet. Gynecology*. **104**, 50–55 (2004).
- Fritz, M. A. & Guo, S. Doubling time of human chorionic gonadotropin (hCG) in early normal pregnancy: relationship to hCG concentration and gestational age. *Fertil. Steril.* **47**, 584–589 (1987).
- Senapati, S. & Barnhart, K. T. Biomarkers for ectopic pregnancy and pregnancy of unknown location. *Fertil. Sterility*. **99**, 1107–1116 (2013).
- Norman, R. J., Menabawey, M., Lowings, C., Buck, R. H. & Chard, T. Relationship between blood and urine concentrations of intact human chorionic gonadotropin and its free subunits in early pregnancy. *Obstet. Gynecology*. **69**, 590–593 (1987).
- Asar, Ö., Ritchie, J., Kalra, P. A. & Diggle, P. J. Joint modelling of repeated measurement and time-to-event data: an introductory tutorial. *Int. J. Epidemiology*. **44**, 334–344 (2015).
- Wulfsohn, M. S. & Tsiatis, A. A. A joint model for survival and longitudinal data measured with error. *Biometrics*. **53**, 330–339 (1997).
- Fitzmaurice, G. M., Laird, N. M. & Ware, J. H. Applied longitudinal analysis 2nd edition. (Wiley, 2004).
- Elandt-Johnson, R. C. & Johnson, N. L. Survival models and data analysis. (John Wiley & Sons 1999).
- Crowther, M. J., Abrams, K. R. & Lambert, P. C. Joint modelling of longitudinal and survival data. *Stata Journal*. **13**, 165–184 (2013).
- Rizopoulos, D. Joint models for longitudinal and time-to-event data: with applications in R. 50–51 (Chapman and Hall/CRC, 2012).
- Marriott, L., Zinaman, M., Abrams, K. R., Crowther, M. J. & Johnson, S. Analysis of urinary human chorionic gonadotrophin concentrations in normal and failing pregnancies using longitudinal, Cox proportional hazards and two-stage modelling. *Ann. Clin. Biochemistry*. **54**, 548–557 (2017).
- Sweeting, M. J. & Thompson, S. G. Joint modelling of longitudinal and time-to-event data with application to predicting abdominal aortic aneurysm growth and rupture. *Biometrical Journal*. **53**, 750–763 (2011).
- Rizopoulos, D. JM: An R package for the joint modelling of longitudinal and time-to-event data. *J. Stat. Software*. **35**, 1–33, <https://doi.org/10.18637/jss.v035.i09> (2010).
- Hickey, G. L., Philipson, P., Jorgensen, A. & Kolamunnage-Dona, R. joineRML: a joint model and software package for time-to-event and multivariate longitudinal outcomes. *BMC Med. Res. Methodol.* **18**, 50 (2018).
- Crowther, M. J. merlin-a unified modelling framework for data analysis and methods development in Stata. Preprint at <https://arxiv.org/abs/1806.01615> [Submitted] (2018).
- Johnson, S. & Marriott, L. hCG levels can decline pre- or post- onset of bleeding in early loss. PO99 @ Fertility 2019 Conference (2019).
- Foo, L. *et al.* Peri-implantation urinary hormone monitoring distinguishes between types of first-trimester spontaneous pregnancy loss. *Paediatric and Perinatal Epidemiology*. **00** (Special Issue: Leveraging Technology), 1–8 (2019).
- Tiplady, S., Jones, G., Campbell, M., Johnson, S. & Ledger, W. Home ovulation tests and stress in women trying to conceive: a randomized controlled trial. *Hum. Reproduction*. **28**, 138–151 (2013).
- Weddell, S. *et al.* Home ovulation test use and stress during subfertility evaluation: subarm of a randomized controlled trial. *Women's Health*. **15**, 1745506519838363 (2019).
- Tewary, S. *et al.* Impact of sitagliptin on endometrial mesenchymal stem-like progenitor cells: a randomised, double-blind placebo-controlled feasibility trial. *E. Bio. Medicine*. **51**, 102597 (2020).
- Coomarasamy, A. *et al.* A Randomized Trial of Progesterone in Women with Recurrent Miscarriages. *N. Engl. J. Medicine*. **373**, 2141–2148 (2015).
- Coomarasamy, A. *et al.* A Randomized Trial of Progesterone in Women with Bleeding in Early Pregnancy. *N. Engl. J. Medicine*. **380**, 1815–1824 (2019).
- Jukic, A. M. Z., Weinberg, C. R., Baird, D. D. & Wilcox, A. J. Lifestyle and reproductive factors associated with follicular phase length. *J. Women's Health*. **16**, 1340–1347 (2007).
- Small, C. M. *et al.* Menstrual cycle characteristics: associations with fertility and spontaneous abortion. *Epidemiology*. **17**, 52–60 (2006).
- Miro, F. *et al.* Sequential classification of endocrine stages during reproductive aging in women: the FREEDOM study. *Menopause*. **12**, 281–290 (2005).
- Gaskins, J. T., Fuentes, C. & De la Cruz, R. A Bayesian Nonparametric Model for Predicting Pregnancy Outcomes Using Longitudinal Profiles. Preprint <https://arxiv.org/abs/1711.01512> (2017).
- De la Cruz, R., Fuentes, C., Meza, C., Lee, D. & Arribas-Gil, A. Predicting pregnancy outcomes using longitudinal information: a penalized splines mixed-effects model approach. *Stat. Medicine*. **36**, 2120–2134 (2017).
- Lukacs, P. M., Burnham, K. P. & Anderson, D. R. Model selection bias and Freedman's paradox. *Ann. Inst. Stat. Mathematics*. **62**, 117–125 (2010).
- Zhang, D., Chen, M., Ibrahim, J. G., Boye, M. E. & Shen, W. JMFit: a SAS macro for joint models of longitudinal and survival data. *Journal of Statistical Software*. **71** (2016).
- Rizopoulos, D. Dynamic Predictions and Prospective Accuracy in Joint Models for Longitudinal and Time-to-Event Data. *Biometrics*. **67**, 819–829 (2011).
- Andrinopoulou, E., Eilers, P. H. C., Takkenberg, J. J. M. & Rizopoulos, D. Improved dynamic predictions from joint models of longitudinal and survival data with time-varying effects using P-splines. *Biometrics*. **74**, 685–693 (2018).
- Ibrahim, J. G., Chu, H. & Chen, L. M. Basic Concepts and Methods for Joint Models of Longitudinal and Survival Data. *J. Clin. Oncology*. **28**, 2796–2801 (2010).
- Johnson, S. *et al.* Development of the first urinary reproductive hormone ranges referenced to independently determined ovulation day. *Clin. Chem. Laboratory Medicine*. **53**, 1099–1108 (2015).

Author contributions

N.B.A. conducted the statistical analysis, drafted and reviewed the manuscript. L.M. processed the data and reviewed the data and manuscript. S.J. developed the sample collection protocol, managed the clinical study and reviewed the data and manuscript. M.J.C. and K.R.A. provided support with statistical analysis and reviewed the manuscript.

Competing interests

M.J.C. and K.R.A. were awarded funding by SPD Development Company Ltd (Bedford, UK) as part of an MRC IMPACT PhD studentship. N.B.A. is the recipient of the MRC IMPACT PhD studentship funded in collaboration with SPD Development Company Ltd. (Bedford, UK). L.M. and S.J. are employees of SPD Development Company Ltd. (Bedford, UK).

Additional information

Correspondence and requests for materials should be addressed to N.B.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020