General enquiries on this form should be made to:
Defra, Science Directorate, Management Support and Finance Team,
Telephone No. 020 7238 1612
E-mail: research.competitions@defra.gsi.gov.uk

SID 5 Research Project Final Report

- **Note**
  In line with the Freedom of Information Act 2000, Defra aims to place the results of its completed research projects in the public domain wherever possible. The SID 5 (Research Project Final Report) is designed to capture the information on the results and outputs of Defra-funded research in a format that is easily publishable through the Defra website. A SID 5 must be completed for all projects.
- This form is in Word format and the boxes may be expanded or reduced, as appropriate.
- **ACCESS TO INFORMATION**
  The information collected on this form will be stored electronically and may be sent to any part of Defra, or to individual researchers or organisations outside Defra for the purposes of reviewing the project. Defra may also disclose the information to any outside organisation acting as an agent authorised by Defra to process final research reports on its behalf. Defra intends to publish this form on its website, unless there are strong reasons not to, which fully comply with exemptions under the Environmental Information Regulations or the Freedom of Information Act 2000.
  Defra may be required to release information, including personal data and commercial information, on request under the Environmental Information Regulations or the Freedom of Information Act 2000. However, Defra will not permit any unwarranted breach of confidentiality or act in contravention of its obligations under the Data Protection Act 1998. Defra or its appointed agents may use the name, address or other details on your form to contact you in connection with occasional customer research aimed at improving the processes through which Defra works with its contractors.

Project identification

1. Defra Project code: SE3040

2. Project title:
   A preliminary analysis of existing data to provide evidence of a genetic basis for resistance of cattle to infection with M. bovis and for reactivity to currently used immunological diagnostic tests.

3. Contractor organisation(s):
   The Roslin Institute, University of Edinburgh
   Scottish Agricultural College, Edinburgh
   Veterinary Laborartories Agency, Weybridge

4. Total Defra project costs (agreed fixed price): £ 147,608

5. Project: start date ..... 01 July 2007
   end date ..... 30 June 2008

6. It is Defra's intention to publish this form.
   Please confirm your agreement to do so.....**YES** ☒ **NO** ☐

   (a) When preparing SID 5s contractors should bear in mind that Defra intends that they be made public. They should be written in a clear and concise manner and represent a full account of the research project which someone not closely associated with the project can follow.

   Defra recognises that in a small minority of cases there may be information, such as intellectual property or commercially confidential data, used in or generated by the research project, which should not be disclosed. In these cases, such information should be detailed in a separate annex (not to be published) so that the SID 5 can be placed in the public domain. Where it is impossible to complete the Final Report without including references to any sensitive or confidential data, the information should be included and section (b) completed. NB: only in exceptional circumstances will Defra expect contractors to give a "No" answer.

   In all cases, reasons for withholding information must be fully in line with exemptions under the Environmental Information Regulations or the Freedom of Information Act 2000.

   (b) If you have answered NO, please explain why the Final report should not be released into public domain

Executive Summary

7. The executive summary must not exceed 2 sides in total of A4 and should be understandable to the intelligent non-scientist. It should cover the main objectives, methods and findings of the research, together with any other significant events and options for new work.

There is both anecdotal evidence pointing to genetic variation for resistance of cattle to infection of M. *bovis*, and published experimental evidence in deer for significant genetic variation in resistance and reactivity to diagnostic tests. However this has not been properly quantified in the cattle population in the UK and it remains a possibility that such genetic variation exists and is a factor influencing the outbreak currently observed in the UK. The genetic variation may be expressed in resistance to infection, in the response to the diagnostic tests, or both. Therefore this study tested these hypotheses using the data collected during the current outbreak on animals that react to the diagnostic test and/or exhibit disease and combining this data with industry databases, particularly dairy databases, that contain additional information on herd mates and pedigree.

To achieve the aims of examining the extent of genetic variation, pedigree information on Holstein Friesian dairy cows was obtained from industry databases on both animals present in the VetNet TB database and their contemporaneous herdmates. This collected information was subject to a quantitative genetic analysis using models that are appropriate to the epidemiological context.

The principal findings from the analyses are:

- Heritable variation in the outcome of TB breakdowns in terms of whether or not a cow is culled by the end of the breakdown. This observation was repeated when the outcome was defined more stringently in requiring 'diseased' animals to be confirmed by observation of lesions or bacteriology. The extent of the genetic variation in the liability for these outcomes was estimated to be 0.16 (s.e. 0.02) and 0.18 (s.e. 0.04) respectively.
- The genetic correlation between the liability of being culled and confirmed with TB and milk yield was -0.48 (s.e. 0.13) suggesting that selection for milk yield has not contributed to the current epidemic and selection for TB resistance would not conflict with the component of the economic breeding goal associated with improving yield of milk (which is still substantial but of diminishing importance).
- The classification of 'Clear', IR, and R from a single test was heritable, in support of the finding above. The heritability on the liability scale for a *single test* was 0.07.
- The heritabilities of skin test values for either M. *avium* or M. *bovis* were low, not detectably distinct from 0. The values were biased due to the lack of skin test values for 'Clear' animals, since selection has been carried out on their difference as part of the culling difference.

The first implication is that within a herd there is heritable variation in individual risk for susceptibility to TB. At a prevalence of 7%, so that 7% of the herd is culled during the breakdown, a value chosen since it approximated the average value found in the dataset, the individuals within the herd have differing probabilities of being culled due to their genotype, with probabilities varying from 1% to 22%. Following on from this, it is clear that genetics can play an important role in control strategies for TB by reducing the incidence of herd breakdowns and reducing the impact of herd breakdowns given their occurrence. It is not anticipated, nor even suggested (since it is assumed that the variation is not complete resistance) that this would be the sole strategy. It is clear that with a disease such as TB a number of co-ordinated strategies need to be undertaken and genetics has the potential to make a continuing substantial contribution as part of this wider effort.

A second implication, of wider relevance than bovine TB, is that government surveillance data can be linked to industry databases to provide valuable evidence on population wide risk factors, including genetics. Furthermore this can be achieved relatively speedily, since these results were obtained within 12 months, from scratch. *One of the deliverables from this project will further speed up this process* in cattle, since cross reference look-up tables have been set up which will facilitate the linking of key herd and animal identification fields across databases in any future study associated with TB or other diseases in cattle.

Finally a plan for implementation and supporting R&D is proposed. Implementation steps vary in lead in times from immediate opportunities to three years.

Project Report to Defra

8. As a guide this report should be no longer than 20 sides of A4. This report is to provide Defra with details of the outputs of the research project for internal purposes; to meet the terms of the contract; and to allow Defra to publish details of the outputs to meet Environmental Information Regulation or Freedom of Information obligations. This short report to Defra does not preclude contractors from also seeking to publish a full, formal scientific report/paper in an appropriate scientific or other journal/publication. Indeed, Defra actively encourages such publications as part of the contract terms. The report to Defra should include:
   - the scientific objectives as set out in the contract;
   - the extent to which the objectives set out in the contract have been met;
   - details of methods used and the results obtained, including statistical analysis (if appropriate);
   - a discussion of the results and their reliability;
   - the main implications of the findings;
   - possible future work; and
   - any action resulting from the research (e.g. IP, Knowledge Transfer).

Background

There is both anecdotal evidence pointing to genetic variation for resistance of cattle to infection of M. *bovis*, and published experimental evidence in deer for significant genetic variation in resistance and reactivity to diagnostic tests. However this has not been properly quantified in the cattle population in the UK and it remains a possibility that such genetic variation exists and is a factor influencing the outbreak currently observed in the UK. The genetic variation may be expressed in resistance to infection, in the response to the diagnostic tests, or both. Therefore this study tested these hypotheses using the data collected during the current outbreak on animals that react to the diagnostic test and/or exhibit disease and combining this data with industry databases, particularly dairy databases, that contain additional information on herd mates and pedigree.

Objectives

To achieve the aims of examining the extent of genetic variation, breed and pedigree information will be obtained from industry databases on animals present in the VetNet TB database, and similarly on contemporaneous herdmates, and this collected information will be subject to a quantitative genetic analysis using models that are appropriate to the epidemiological context.

Specific objectives are:
1. Identify herds and cattle present in the Defra VetNet TB database in CTS and dairy industry databases.
2. Using the linkage established between VetNet, CTS and industry databases, develop an integrated database identifying both animals appearing in the VetNet TB database and their contemporaries present at the time of testing with pedigree and relevant aspects of performance.
3. Define and calculate a set of epidemiological and genetic covariates to be used for modelling.
4. Construct and refine models of TB-related data accounting for genetic, operational and environmental factors.
5. Interpret outcomes and develop recommendations from final models.

A glossary of important terms is given in Annexe 1. Items in blue text are described within the Glossary.

The contractors believe these objectives have been met in full, and in some aspects exceed what might be expected from a preliminary analysis, both in deliverables and in interpretation.

Overview of Study Methods.

Prior to detailing the Materials & Methods used in the analysis to a depth appropriate for a scientific review, a brief summary will be given to outline the process carried out and the issues that arise in interpreting such a study. These issues will be addressed in detail throughout the Report.

The data required to carry out this study existed before the project started, however key elements of the data existed on different databases, some owned by Government and its agencies, and some

owned by industry organisations. Therefore the first task was to link these databases together so that the dispersed information could be brought together in an ordered fashion. This was considered to be achievable but the degree of success was unknown at the outset.

The data being linked and extracted was field data collected as part of routine surveillance for *M. bovis* infection in herds. Whilst this has an important strength of relevance, external validity and, potentially, of data volume, it has the weakness that it is obtained through the diversity of circumstances occurring in the field, which are not subject to the same level of control that would occur in an experimental study. This requires validation of the data to ensure that it is coherent and appropriate for the analysis to be undertaken, since poor quality data leads to imprecise and possibly misleading results – 'rubbish in, rubbish out' very much applies to data analyses. In this study a major challenge was that such large scale analysis of field data for disease is rare and the soundness of both the data and the interpretation must be judged from both a genetic and an epidemiological perspective. Examples of epidemiological issues addressed below are exposure, and whether or not the testing procedure used for diagnosis identifies all animal that are diseased within a herd and/or wrongly identifies some animals as diseased.

From a modelling perspective, the approach described below avoids an analysis of factors influencing why one herd suffers from TB and another does not. Rather, it seeks to examine the cows within a herd that become diseased and those cows that do not, and analyses the genetic variation in particular that may lie beneath this occurrence. The information obtained within herds is then pooled over many herds to obtain precision. A key source of variation in livestock studies for estimating the genetic variance is the variance between sires, and, typically, within a herd, it is possible to obtain such information since more than one sire is used to generate a cohort of animals.

One issue of pooling variation over many herds is that each herd will potentially have a different prevalence, and it is intuitive that the degree to which genetics may be expressed might depend on prevalence, for example when a disease is very rare those at moderate genetic risk are unlikely to be distinguishable from those with low genetic risk since both groups are highly unlikely to become diseased. An established modelling approach, adopted here, is to consider the risk on an underlying continuous scale of risk called liability and to relate the genetic variance observed within a herd to the liability scale using its observed prevalence (see Glossary).

Materials and Methods

Tuberculin testing in cattle within GB, using skin tests. The data used in this study were the results of skin tests and bacteriological tests carried out on dairy cattle herds during a TB incident (breakdown) within the herd, occurring in GB from 2000 to 2007. As described in detail below the actual data available for analysis is formed by the procedures of the compulsory skin testing programme using single intradermal comparative cervical tuberculin skin test to detect TB in cattle within GB. The salient points of the testing regime in relation to this study are as follows. The basic test given to the cows follows the procedures described by Lesslie and Herbert (1975) involving injection with M. *avium* and M. *bovis* with diagnosis depending on the differential skin thickening in response to these inoculations. During the period covered in this dataset there was no use of the blood based γ-interferon test in routine herd tests for TB, although its use in parallel or in series with the skin test became mandatory in specified circumstances in October 2006. The precise interpretation of the tests is a standard procedure applied nationally across the GB (de la Rua-Domenech et al., 2006), and this varies slightly from the paper of Lesslie and Herbert (1975). Following a single test, an animal is classified as: (i) a non-reactor ('Clear') in which the response to M. *bovis* is judged in a defined way to be less than or equal to that of M. *avium*, and hence indicative of freedom from bovine TB; (ii) inconclusive reactors (IR) where the M. *bovis* reaction exceeds that of M. *avium* by up to 4mm; (iii) reactors (R) in which the response is deemed to be strongly indicative of disease (i.e. *M. bovis* reaction more than 4mm greater than *M. avium* reaction). Discovery of R and/or IR animals within a herd prompts further tests. All R animals are culled immediately, and any animals that are classified as IR on three consecutive tests are also culled. The herd is placed under movement restrictions and testing at intervals of 60 days continues until one or two (depending on post-mortem findings) wholly negative herd tests are obtained, when the breakdown is formally 'closed' and movement restrictions are lifted. Routine surveillance for TB in herds is conducted nationally at 1,2, 3 or 4 yearly testing intervals that vary among geographical regions reflecting the perceived risk of disease in the region. This on-farm

tuberculin herd testing regime is supplemented by “passive” TB slaughterhouse surveillance of cattle carcases during normal meat production.

Databases and data extraction. Data was extracted from 2 complementary but previously unlinked databases. The first was a database compiled by the VLA (Weybridge Laboratory) from the Animal Health disease management database, VetNet, which maintains a record of tuberculin testing associated with herd breakdowns in GB. VetNet records details about the herd e.g. identity, defined by geographical location (county, parish, holding; CPH); the breakdown, e.g. initial date and confirmed end date; skin testing including the number of animals tested on that date, and the ear tag number of all cattle found to have tested ‘IR’ or ‘R’ together with the M. avium and M. bovis reaction differentials leading to this categorisation. Two points should be noted: the ear tag number identifies a bovine animal uniquely, conforming to EU regulation; and, a serious deficiency, animals testing ‘Clear’ at a skin test are not recorded on VetNet. Consequently there is no information contained within the database on the identity of these animals (eartag) and their skin tests results and hence the distribution of *M. avium* and *M. bovis* scores resulting in an ‘Clear’ test. A caveat to this statement is that negative skin test results are recorded within VetNet for some Dangerous Contacts, i.e. IRs and ‘Clear’ animals deemed to have been exposed to *M. bovis* and removed at the discretion of the Veterinary Officer, typically only in TB breakdowns where there is PM evidence of infection.

The second database was compiled from commercial companies recording performance in dairy herds. This database resulted from data extracted from two sources NMR plc (Chippenham) and the CDI performance database managed by Holstein UK (Scotsbridge House, Rickmansworth). Each of the databases contained records on all milking cows within a herd present on a performance test. These milk recording tests are carried out at intervals of 28 days or more (depending on the commercial arrangement with the farmer). For each cow recorded the following information is available: pedigree in the form of a nominated sire and dam; date of birth; and various measures of performance on each test date such as milk yield. The recording of the unique UK ear tag number used by VetNet increases over time, although in neither database is this the primary key identifying a cow.

The databases were linked using the unique ear tag, but since this was not the primary key of the industry databases its value, where recorded, was of variable quality, did not always conform to the strict tag format and was of an inconsistent nature even within farm. This necessitated the use of rule-based but inexact matching procedures, although any introduced errors of pedigree is likely to have reduced estimates of heritability, and this is addressed in more detail in the Discussion. The procedures involved re-formatting the ear tag field in both databases to make them standardised in both cases. This was followed by a number of steps and subsequent matches: 1) removing all non-numeric characters and spaces; 2) extracting the herd part of the ear tag from the animal part; 3) extracting the herd and animal parts of the ear tag and then removing leading zeros from each; 4) converting ‘O’ and ‘o’ to zero. A unique key was added to the destination table and animals added to that table based on descending degrees of match starting from a) exact match on whole unformatted ear tag; b) exact match on numeric version of ear tag; c) exact match of herd plus animal parts separate (but only where each part was at least 4 characters long); d) exact match on herd and animal with no leading zeroes (again only where each part was at least 4 characters long). If an animal was found at a high degree of match then it was not added a second time at lower degrees of match. A summary of this matching process is given in Annexe 2.

Given a matching between an animal in the VetNet database and the industry databases on a performance test day close to skin test date it was possible to infer the contemporary animals that were present at the skin test but were categorised ‘Clear’ (and were therefore not in the VetNet database). For all animals identified either as being VetNet or as a contemporary, the performance test date closest to the skin test date was extracted. A check on the quality of this inference was carried out by comparing the estimated number of cows tested identified from the industry database and the number recorded within the VetNet database.

Data editing. The initial matching process identified in excess of 450,000 records on animals contained on VetNet and their contemporaries, although included more than 1 record per cow. A number of data edits were made to improve quality assurance on the records to be analysed and these included edits made for epidemiological reasons. Initial edits were on information quality: a further more stringent review was made of herd matching and test dates; all cattle other than Holstein Friesian cows were

removed (Holstein Friesian cows are the common black and white dairy cows, and their gene pool comprises 95% of all dairy cows in the UK); and herds with less than 5 cows were removed. The following editing of breakdowns was made for epidemiological reasons.

1. Only the first breakdown in a herd, judged from within the extracted dataset, was included, with the intention that possible complications of changed immunological status arising from prior breakdowns were minimised. It is established that immune reactions change as a consequence of prior exposure to a pathogen: for example it is possible to acquire immunity from exposure to a pathogen or a related pathogen (for example, in humans, cowpox protecting against smallpox). The elucidation of these changes for TB were beyond the objectives of this study.
2. Breakdowns occurring earlier than 2000 were removed i.e. all records associated with that breakdown: since the dataset screened started in 1995, all herd breakdowns included in the analysis were from herds that had been clear for 5 years, and that this had been confirmed by at least one herd test, since the maximum length of the testing cycle in the GB is 4 years.
3. Breakdowns that were not closed after 2 years from onset were removed to limit the possibilities of including repeated breakdowns within a herd that were not recognised as such within VetNet.
4. Breakdowns with less than 2 'R' cows were removed, with the intention of ensuring that the breakdown was not due to single, possibly imported, cow.
5. For cows appearing in more than one breakdown, records apart from the first were removed from the data, for the same reason as item 1.

One possibility that was addressed was that first lactation animals may have been subject to a different exposure risk due to differential management over the course of a breakdown compared to later lactations. In the extreme case, exposure to the source of infection, and opportunity for transmission may only have occurred in one of the groups. Given the underlying principle was to identify groups in which individuals within a group had *a priori* equal opportunities for infection, although the opportunities may differ between groups. The following edits were made to maintain this principle:

6. If at least 90% of 'R' cows (class 7, Table 1) occurred within a single age group then all other groups were deleted, since the TB breakdown was deemed to appear as if it originated from a point source within a cohort.
7. If no 1st lactation cows were identified as 'R' (class 7, Table 1) within the breakdown then the entire first lactation cohort was deleted, since the epidemic appeared as if it originated as a propagating epidemic within the milking herd.

In the Results that follow, all TB breakdowns were initially considered, whether or not they had been "confirmed" by lesions at PM or by positive culture. However in the course of the analyses the subset of "confirmed" breakdowns was used for further refinement of analyses. TB breakdowns triggered by cases detected at meat inspection were included in the final dataset.

Following this data editing, records remained on 68,497 cows (Table 1) in 818 herds, with an average of ~80 cows per herd. Pedigrees for up to 5 generations were then extracted using the HUK pedigree database and other available sources, following both sires and dams wherever possible, for all remaining cows with breakdown data or inferred breakdown data (i.e. 'Clear'), resulting in a total of 228,508 animals being included in the pedigree file.

Table 1. Numbers of cows in each of 7 categories defined by the passage of the cow through the breakdown, including the scoring of the cow in ultimate fate models.

Class	Ultimate Fate C & D	Stage 1	Stage 2	Stage 3	Number	%
1	0	Clear			58,086	84.8
2	0	IR	Clear		4,444	6.5
3	0	IR	IR	Clear	986	1.4
4	1	IR	IR	IR	183	0.3
5	1	IR	IR	R	118	0.2
6	1	IR	R		410	0.6
7	1	R			4,270	6.2
Total					**68,497**	

Models fitted. TB epidemiologists are careful to differentiate the skin test as an imperfect means of

identifying animals with disease, and the disease itself. Therefore a range of models were fitted to explore these different aspects, in part because it was necessary to do so to achieve our objectives. In this context it is natural to start with models for the skin tests since, in this study, disease is identified via skin tests, although not solely by skin tests.

(i) A priori, the individual skin test values recorded in VetNet have a strength in that it is continuous data but it has a considerable weakness in that it is available only on those cows that are 'IR' or 'R' and results in a bias. There is value in the models, since there are data methods and approximations that may be used to correct such biases and the estimates obtained are suggestive. For example, if it was found that skin test values in this biased sample were of moderate heritability this could not be ignored when interpreting any skin test procedure. This model is termed the Skin Test Model. However, the data selection is severe, and interpretation is complicated by unestimable responses to disease.

(ii) One way of overcoming the bias from the selection of data, at a cost of losing the continuous data, is to replace the test value with the progressive status 'Clear', 'IR' and 'R'. The bias is removed as all cows present can be included in the analysis, since cows that were not included in VetNet are 'Clear'. However the data now consists of a series of test outcomes for each cow, since a cow is tested several times for a breakdown. Although set against a background of an epidemic, it might be postulated that cows will exhibit a 'characteristic' response to the skin test. This is most directly modelled by the Herd Test-Date Model where all 'Clear' tests for a cow were included. The missing 'Clear' tests are identified by a Herd Test-Date appearing in VetNet, but the cow, although it is expected to be present, is not identified in VetNet as an 'IR' or 'R'. Such a model could detect, for example, whether 'IR' appears randomly among the herd or are an attribute of the cow or its genetics. However such a model remains focussed on skin test response, rather than as the occurrence of a disease, and looks at the results as repeated measures under stable conditions, thus would be very suitable for modelling the data in disease free herds. In interpreting the results as liability to TB infection there is an underlying assumption the liability is the responsiveness to skin test. This is a questionable assumption.

(iii) A variation on this approach which does go towards recognising the progression of a disease over time, is to consider an 'IR' as a step towards possible culling. This approach was modelled by a Continuation Ratio Model, which is an established model of progression between stages (for example in educational progression). The Continuation Ratio Model is also favoured by epidemiologists as a suitable method of dealing with a process which comprises ordered responses. It is appropriate when the process represents a progression of stages (in this case, each stage is a test) so that individuals must pass through each lower stage before progressing to the next. Outcome at each stage is represented by pass (i.e. 'Clear') or fail (i.e. not 'Clear', which could be either an 'R' or an 'IR') and modelled using an appropriate function for binary traits. The weakness of such a model is that, like the Herd Test-Date Model, the skin test responsiveness itself is taken to be the liability of becoming diseased.

(iv) Therefore further models moved away from using the individual skin tests, and instead looked at outcomes over the whole breakdown period. These were termed Ultimate Fate Models. Philosophically, it is more straightforward and pragmatic than (ii) and (iii): there is a testing procedure which identifies animals that are deemed to have disease and are consequently culled, whilst those not culled are deemed to be free of disease. This is a classical epidemiological testing framework, parameterised by the sensitivity (Se) of the test and the specificity (Sp) of the test. Tests of different Sp and Se may have different properties. The simplest such model was Ultimate Fate 'Model C' (C for culling) where animals were only classified by whether or not a cow was culled during the breakdown. The Ultimate Fate 'Model D' (D for Disease) classified by whether the cow was culled and was subsequently confirmed to have disease. The epidemiological distinction is that the first has incomplete specificity and the second has complete specificity.

Skin Test Model. A bivariate animal model was fitted to the skin test scores treating M. *avium* response and M. *bovis* response as separate but correlated traits. Only first test scores of cows judged 'R' or 'IR' at first test could be included in the analysis (10,411 records from Classes 2 to 7 in Table 1) to reduce the impact of bias since, within the herd breakdown and given the limitations on data availability, these will be the least affected by disease. As the distributions of both M. *avium* and M. *bovis* response are highly skewed, the traits were transformed using natural logarithms. This transformation helped to make the distributions appear more normally distributed and make the analyses conducted more robust, since in a very skewed distribution a small subset of the data may be given considerable weight in the analyses. The fixed effects included were herd breakdown, lactation group (2 classes: 1st lactation, or 2nd and later lactations grouped together), the interaction between herd breakdown and lactation group, and month of test. Age (in days, at the start of the breakdown) and milk yield (at the

test day closest to the start of the breakdown), were included in the model as covariates, both with quadratic relationships between the covariate and the skin test data. Genetic effects for each cow were fitted as random effects modelled using their pedigree i.e. a full animal model was fitted to the data. As stated in the description of the testing procedure earlier, these are selected on the difference between the *bovis* and *avium* scores. Correction for this selection was complicated by the fact that whilst selection was on the difference on the observed scale, the analysis was on the log scale. This is returned to in the Discussion.

Herd Test-Date Model. These models recognise the full testing sequence. Therefore a 'Clear' cow in Class 1 of Table 1 is recognised as the outcome of a sequence of tests during a breakdown, with all test results being 'Clear'. Since the herd is closed from the start of the breakdown it is possible to infer that, for example, a cow that is not present in the VetNet database would have tested 'Clear' at every single skin test carried out during the breakdown. Therefore using the following procedure an extended data set could be developed by the following rules:

(a) Animals present at the start of the herd breakdown were always present at each of the subsequent tests unless culled as an 'R' or a 3-fold 'IR'.

(b) Unless indicated otherwise by the VetNet database, all cows present at a given test-date were assumed to have given a 'Clear' outcome.

The trait analysed was a binary variable with 'Clear' = 0, otherwise 1. The model was applied to the extended data set and effects were estimated on the liability scale using a complementary log-log link function. The main contemporary grouping was breakdown by test-date, and age at each specific test in days was included in the model as a quadratic polynomial. Genetic and permanent environmental effects for each cow were fitted as random effects, with genetic effects modelled using their pedigree. Given that test data over a maximum of two years had been inferred, neither lactation number nor milk yield was available. Note also that, as young stock of less than 12 months of age are not routinely tested, no data was inferred for cows less than 12 months old.

Continuation Ratio Model. This model was constructed within a continuation ratio framework (see Agresti, 2002, page 289) which is particularly appropriate for staged data where the response is discrete. A breakdown can be viewed as having up to 3 stages in which a cow progresses towards being culled (see Table 1). For each of the three stages there are two 'hurdles' defined by not being 'Clear' and, if not a 'Clear', becoming an 'R'. Continuation from Stage 1 to test Stage 2 and from Stage 2 to Stage 3 was assumed to occur only for 'IR' cows. This progression through the breakdown defined seven classes of cow as shown in Table 1, defined by the number of 'IR' tests a cow receives.

Models fitted to the binary trait estimated effects on the liability scale using a complementary log-log link function. Effects in the model were supplemented by a fixed effect with six levels, representing the six 'hurdles', two for each stage, together with genetic and permanent environmental effects for each cow fitted as random effects, with genetic effects modelled using their pedigree. This model was applied to the full data set.

Ultimate Fate 'Model C'. The data from the series of skin tests was simply summarised by defining the ultimate fate (u_C) of the cow as a consequence of the testing during the breakdown: $u_C = 1$ if a cow has been culled at the completion of the breakdown irrespective of whether it was culled as an 'R' or as a multiple 'IR'; and $u_C = 0$ otherwise (Table 1). Models fitted to this binomial variable estimated effects on a liability scale using a complementary log-log link function. The consequence of this model is that estimates of heritability and other variance components are, in principle, independent of the prevalence within the different herd breakdowns. As first lactation animals may have experienced a different exposure risk compared to later lactations, cows were grouped into milking heifers and later lactations. The fixed effects included were herd breakdown, lactation group (2 classes: 1st lactation, or 2nd and later lactations grouped together), interaction between herd breakdown and lactation group, and month of test. Age (in days, at the start of the breakdown) and milk yield (at the test day closest to the start of the breakdown), were included in the model as covariates, both with quadratic relationships between the covariate and liability. Genetic effects for each cow were fitted as random effects modelled using their pedigree i.e. a full animal model was fitted to the data. Fitting the interaction between herd breakdown and lactation group identifies both lactation groups as potentially having different opportunities for becoming diseased over the period of the breakdown.

In addition a bivariate analysis was carried out including test day milk yield closest to the start of the breakdown as a separate but correlated trait. The bivariate analysis allows both genetic and phenotypic correlations between the two traits to be estimated. The model for milk yield included age at the milk recording and stage of lactation in days, both modelled as quadratic polynomials.

Farm management practices, weather conditions, TB spoligotypes, and force of infection vary across regions of Great Britain, resulting in environmental and potential epidemiological differences between regions. Our models assume that the ratio of the genetic variance to the environmental variance is constant on our underlying liability scale. If environmental differences exist they would be expressed as variation in heritability across regions. To check for any differences in heritability, Ultimate Fate 'Model C' was extended to analyse regions North, West and Wales, fitting a common covariance between regions. The East and Scottish regions were ignored due to limited amounts of data.

Ultimate Fate 'Model D'. This model was very similar to the ultimate fate model C, except that u was defined not only by the culling following the skin test outcomes, but also on whether or not bacteriological or slaughterhouse evidence following culling supported the diagnosis of TB. Therefore in 'Model D', $u_D = 1$ only if $u_C = 1$ *and* either M. *bovis* was found in bacteriological cultures following culling or evidence of lesions were found at slaughter, otherwise $u_D = 0$. This entailed additional data editing to remove whole breakdowns and whole cohorts that did not conform to the guidelines given above, and resulted in records on 18,339 cows in 221 herds. A total of 76,284 animals were included in the pedigree. The effects fitted were as for Model C.

Results

Table 1 gives the number and percentage of observations for each of the 7 classes used in the Continuation Ratio Model. To simplify the presentation of results where covariate adjustment has been used, we define an average cow as a cow which is 1600 days old and/or yields 25 kg of milk per day.

Ultimate Fate Models C & D. Results from these models were similar. All effects in both models removed a significant amount of variation. From Model C, the probability of an average cow being culled as an R or a multiple IR was highest in June and July and lowest in November and December. The difference in culling between summer and winter months was statistically significant ($P < 0.05$). The solutions for month of test from Model D differed slightly to those from Model C in that the probability of culling was highest in February and lowest in October ($P<0.05$). Figure 1 shows that the probability of an cow of 1600 days of age being culled as a reactor decreases as her daily milk yield increases, falling from 0.086 to 0.042 (Model C) and 0.072 to 0.025 (Model D). Figure 1 also shows how the probability of being culled varies depending on the age of the cow. The probability increases from first calving to around 6 years of age, then decreases although the s.e. is also increasing over this later range. Heritability of liability was estimated as 0.16 ± 0.02 from Model C and 0.18 ± 0.04 from Model D (Table 2), the smaller data set leading to a loss in precision; nevertheless, for both models, the evidence for genetic variation was highly significant ($P<0.001$).

Figure 1. The effect of daily milk yield (left) and age (right) at the start of the breakdown on the risk of being culled over the course of the breakdown.

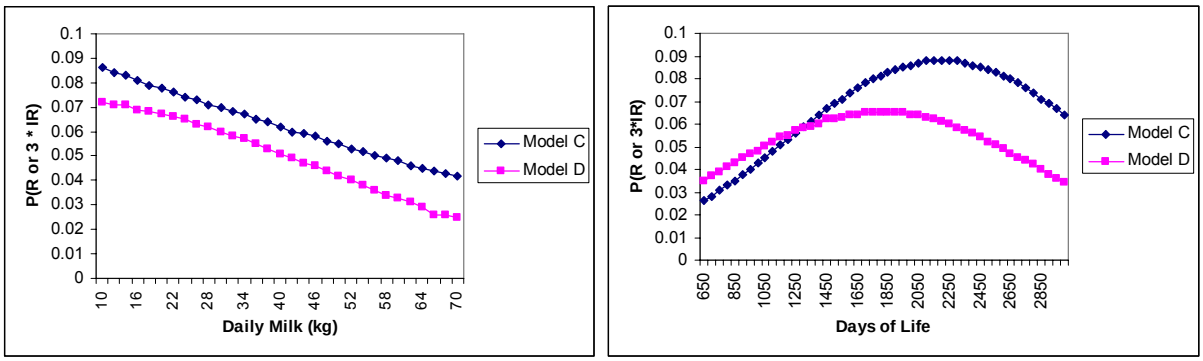


Table 2. Estimates of heritability and the fraction of the variance due to pemanent environment (c^2) for Ultimate Fate models, Continuation Ratio (CRM) and Herd Test-Date (HTD) models. Standard errors are in parentheses and N/A indicates parameter was not part of the model.

Model	Ultimate Fate 'C'	Ultimate Fate 'D'	CRM	HTD
h^2	0.155 (0.019)	0.181 (0.044)	0.077 (0.012)	0.073 (0.013)
c^2	N/A	N/A	0.159 (0.013)	0.329 (0.013)

Results from the bivariate analysis of ultimate fate and daily milk yield of cows whose skin test results were confirmed by either bacteriological or slaughterhouse evidence is given in Table 3. Heritability of liability (0.19) is, as expected, consistent with the estimate from the univariate analysis of 0.18 (Model D). Heritability of milk yield (0.17) is lower than the figure used in national evaluations (~0.50; Egenes, Scottish Agricultural College) possibly due to different methods of estimation (a simple linear model here compared to the more demanding and complex 'random regression test day model'). Both phenotypic and genetic correlations are negative, indicating that both management and genetics contribute to higher yielding cows being less likely to be culled as reactors. The magnitude of the phenotypic correlation is small (-0.07) but that for the genetic correlation is moderate (-0.48). Whilst the expressed phenotype for milk will vary with time over the course of the breakdown through the milking cycle complicating the interpretation, the breeding value for yield as estimated from the model is not variable and has no such complications.

Table 3. Genetic (above the diagonal from top left to bottom right) and phenotypic (below the diagonal) correlations between the risk of being culled as a reactor with confirmation (u_D) and daily milk production. Heritabilities are on the diagonal. Standard errors are given in parentheses.

	u_D	Milk yield
u_D	0.19 (0.042)	-0.48 (0.130)
Milk yield	-0.07 (0.028)	0.17 (0.019)

Estimates of heritability of liability of u_C from thr regions are presented in Table 4. The number of records varied from 9,057 in the North to 45,826 in the West. Heritabilities ranged from 0.19 (West) to 0.48 (Wales) with large s.e.'s arising form the smaller numbers in each data subset. Heritabilities were consistent across the West and North regions, but the heritability of liability in Wales appeared significantly higher (P<0.05). Nevertheless there was no evidence from the regional analyses that large differences existed between the effects of individual sires across the different regions.

Table 4. Regional variation in estimates of heritability of liability based upon the Ultimate Fate Model C. Standard errors are in parentheses.

	West	North	Wales
Number of records	45,826	9,057	11,088
Proportion culled	0.071	0.080	0.077
Heritability	0.19 (0.045)	0.23 (0.12)	0.48 (0.11)

Continuation Ratio Model. All effects fitted in this model were statistically significant. Table 5 gives the probability of a clear result ('Clear'), the probability of an inconclusive result and the probability of being classed a reactor at each of the three possible tests, for a typical first, third and fifth lactation cow. The table shows clearly that the probability of 'Clear' decreases as an animal ages from 1st to 5th lactation, whereas the probability of being a reactor increases with age. The heritability of liability from this model was estimated as 0.08 ± 0.01 (P<0.001), as shown in Table 2. Note that in these models the heritability is for the *outcome of a single stage, not the outcome of the process*, and this will contribute to the lower heritabilities involved.

Herd Test-Date Model. As in the other models, days of life removed a significant amount of variation in the trait. The genetic variance and the cow's permanent environmental variance were both significantly greater than zero. Heritability of liability was estimated as 0.07 ± 0.01 (Table 2) and was statistically significant (P<0.001). As for the Continuation Ratio Model the heritability is lower than the Ultimate Fate models since the heritability is for a single test classification not for a whole process of diagnosis. The total variation attributable to an individual is the sum of the heritability h^2 and c^2, which as shown in Table 2 was ~0.4.

Table 5. For 3 categories of cows, the probability of a non-reactor, P(Clear), an inconclusive reactor, P(IR), and a reactor, P(R), at each of 3 stages.

	P(Clear)	P(IR)	P(R)
Lactation 1 animal 800 days old producing 25 kg milk per day			
Stage 1	0.915	0.073	0.012
Stage 2	0.898	0.096	0.006
Stage 3	0.938	0.057	0.005
Lactation 3 animal, 1600 days old, producing 25 kg milk per day			
Stage 1	0.826	0.131	0.043
Stage 2	0.804	0.173	0.023
Stage 3	0.877	0.102	0.021
Lactation 5 animal, 2400 days old, producing 25 kg milk per day			
Stage 1	0.806	0.141	0.053
Stage 2	0.781	0.191	0.028
Stage 3	0.862	0.112	0.026

Skin Test Model. The estimates for genetic variances for both M. *avium* and M. *bovis* were not significantly different from zero (P>0.05) and heritabilities were low (0.03 and 0.02 respectively). Phenotypically, the traits were correlated (0.34, s.e. 0.010), but no significant genetic correlation was detected, the latter unsurprising given the difficulty of establishing a genetic variance.

Discussion

The project has addressed the objectives and achieved the following items.

- The linkage of the VetNet and industry databases to produce a database containing the necessary genetic and performance information of sufficient size and power for detection of genetic variance in susceptibility to bovine TB. This achieves Objectives 1 and 2.
- The construction of an epidemiologically meaningful set of observations and covariates for the genetic analysis. This achieves Objective 3.
- Extensive analyses of the dataset to determine the extent of genetic susceptibility to TB in dairy cattle, and to test the robustness of the models across a range of issues, discussed in detail below. This achieves Objective 4.
- Considered the implications of these results for management of TB, implementation of genetic strategies and future R&D needs (see below). This achieves Objective 5.

The principal findings from the analyses are:

- Heritable variation in the outcome of TB breakdowns in terms of whether or not a cow is culled by the end of the breakdown. This observation was repeated when the outcome was defined more stringently in requiring 'diseased' animals to be confirmed by observation of lesions or bacteriology. The extent of the genetic variation in the liability for these outcomes was estimated to be 0.16 and 0.18 respectively, both highly significant statistically.
- The genetic correlation between the liability of being culled and confirmed with TB and milk yield was -0.48 (s.e. 0.13) suggesting that selection for milk yield has not contributed to the current epidemic and selection for TB resistance would not conflict with the component of the economic breeding goal associated with improving yield of milk (which is still substantial but of diminishing importance).
- The classification of 'Clear', IR, and R from a single test was heritable, in support of the finding above. The heritability on the liability scale for a *single test* was 0.07.
- The heritability of skin test values for either M. *avium* or M. *bovis* were low, not detectably distinct from 0. The values were biased due to the lack of skin test values for 'Clear' animals, since selection has been carried out on their difference as part of the culling difference.

A number of issues can be identified that may interfere with the estimation of heritability, however in all these cases analysis of theory leads to the conclusion that each issue will lead to underestimation of the true heritability, and the true heritability for TB susceptibility will be greater than has been estimated here. These will be considered in turn.

(i) The use of pedigree in recording databases that is only nominated by the farmer rather than established or confirmed by DNA. However these pedigrees are used for the genetic evaluations of the primary performance traits and these evaluations are well established as being robust. A subset of the pedigrees in the Holstein UK breed society are subject to a quality assurance scheme. Further Woolliams (2006) when examining the impact of wrong or missing sire information, showed that, in this case, wrong pedigree is twice as potent as missing pedigree in *reducing* the accuracy of genetic evaluations. Therefore it is reasonable to conclude that any significant extent of pedigree errors contained within the data will lead to an underestimate of the true heritability of liability.

(ii) An argument can be made that exposure to the pathogen is unequal among the cattle in the herd, or at least cannot be assured to be equal, and this will influence the heritability observed. However the primary source of genetic variation comes from the variation among sire families and the observation of any sire variation suggests that genetic variation is present unless, in some way, the offspring of *individual* sires are managed differently from their contemporaries to promote or repress the opportunity for infection. This seems unlikely as such blatant difference in management would have undermined all well-established genetic evaluations. Therefore rather than equal exposure the cows within a contemporary group have equal opportunity for infection. Nevertheless a valid question is to what degree is heritability affected by incomplete exposure. Heuristically, consider one sub-group of cows within a herd exposed and a sub-group that remain unexposed: in the first group the genetic variance for susceptibility is expressed, whilst in the second group no genetic variation is expressed, therefore genetic variation is diluted. Assuming that exposure is independent of sire family as argued above, the variance between sire families is reduced by this dilution, reducing the heritability. This heuristic explanation can be derived analytically (Bishop and Woolliams, manuscript in preparation).

(iii) A further argument that may be advanced against the heritability observed truly reflecting TB susceptibility is that the measure used is based on a test that is not specific, in that some non-diseased animals are culled, and potentially not sensitive in that some animals that are diseased are not culled. Consider a test that is not specific, and assume that it perfectly sensitive for the present: a consequence will be that the expression of the true variation among sires will be diluted by more cows from more resistant sire families being counted among the 'diseased'. A second consequence will be that the observed prevalence will increase due to these additional cows being counted as 'diseased'. This has a two-fold impact, firstly the variation is reduced and secondly the conversion factor from the observed to the liability scale is reduced due to the increased prevalence. Therefore the estimate of heritability on the liability scale will be open to severe underestimation. In contrast, consider the case where the test is specific in that all animals with a positive test are truly infected, but the test may not be perfectly sensitive as in the case with u_D. Again the sire variation is reduced since the susceptible families are observed as being less extreme than they are. Further the observed prevalence is reduced. Consequently the sire variance is reduced but the conversion from the observed to liability scale is greater than it should be, counteracting in part the reduced sire variance. Theoretical exploration also shows that the net result will be an underestimate, but not to such a severe degree as when the test is sensitive but not specific. These heuristic explanations can be derived analytically (Bishop and Woolliams, manuscript in preparation). A conclusion from this is that estimates of heritability from both u_C and u_D are underestimates but the bias may be more serious in u_C than u_D, because the former is neither perfectly specific nor perfectly sensitive, whilst the latter is perfectly specific but not perfectly sensitive. Therefore, the estimate of u_D is open to less bias than u_C despite a greater standard error. With this perspective and using reasonable estimates for the sensitivity for the test of disease requiring both culling and confirmation through lesions or bacteriology, and correcting for the likely bias, the best estimate of h^2 on the liability scale is in excess of 0.20.

In conclusion, the above supports the general argument that the potential problems listed above will lead to an underestimate of the true h^2 for susceptibility.

One aspect of the concerns about the imperfect specificity and sensitivity of the test procedure based upon skin responsiveness (e.g. heritability of u_C) is that the heritability of the test is in truth the heritability of responsiveness, and that any selection on the outcome of the testing process will merely reduce the skin responsiveness and diminish the existing value of the testing procedure. The term responsiveness could involve one or both of two different things:

- The 'skin-test responsiveness' may be *a property of an individual irrespective of disease*, the type that might be examined within the VEBUS database where field tests of individuals are recorded in

disease free herds. A putative source of such a responsiveness may be an allergy to the skin-test in a differential way so as to interfere with test outcomes. Such a response might be expected to affect both the epidemiological sensitivity (Se) and specificity (Sp), so that an individual may have an immune system that reacts in such a way as to increase its chance of being seen as as a reactor.

- Alternatively the 'skin-test responsiveness' may be regarded as the *probability the test will be positive given that the disease is present*, so that it is seen as an interaction of the individual with the pathogen, and not as a property of an individual alone.

For the first of these possibilities, concerning responsiveness independent of disease, is not supported by the study. The most direct argument is that if the heritable variation were primarily due to skin responsiveness rather than infection, then the change in trait from u_C to u_D reduces the variation between sires in skin responsiveness since offspring that were responsive but in fact had no infection are then grouped with non-responders, and if the sire variation is reduced (the primary source of genetic information) so would the observed heritability on the liability scale. However this clear prediction was not observed, the reverse was observed, consistent with the idea that genetic variation was due to susceptibility due to infection (as described above) giving a clear indication that the genetic variation is indeed linked to the infection. Further, notwithstanding the concerns over the quality of the skin test data, it would have been anticipated that greater evidence for a heritable response would have been observed (see below for further discussion on this). The second of these possibilities is addressed in Annexe 3 where it is firmly argued that the relatively sizeable magnitude of the heritability observed is incompatible with it being caused by responsiveness given accpetable ranges for the sensitivity of the test and the prevalence of TB.

Whilst this is an important observation there are two barriers to further analysis of this aspect. Firstly the field values for specificity and sensitivity of the test appear to have serious noise, and although they are high, there is debate over how high. Reliable information on these values would not only inform control strategies based upon phenotype but would also provide lower limits to possible genetic correlations between skin responsiveness and susceptibility to infection. Secondly, no large scale genetic analysis has been possible on skin responsiveness of cows that are almost certain to be free from infection. This is a major gap that seriously limits the ability to draw firm inferences from data, and so invokes avoidable doubts over control strategies with consequent loss of direction and momentum.

Nevertheless it was part of the study objectives to carry through such an analysis and this was done. The only possible estimates of heritability from for skin responsiveness *per se* from VetNet data were biased since data had come from cows that were either 'IR' or 'R': from Table 1 it is seen that for the most part 'IR' cows are deemed free of infection at the end of the breakdown. Nevertheless the data available were those that had been truncated in relation to the difference between M. *bovis* and M. *avium* response. The empirical outcome was that no significant variation was detectable. However based on the empirical estimates, an estimate of unbiased heritability of the differential responsiveness in the absence of disease can be obtained with some assumptions. Assume the correlation between the responsiveness to the 2 inoculants is zero, which will be conservative if a prior belief is that the true genetic correlation is positive and greater than the phenotypic correlation. Then the heritability of the difference in log responsiveness in the truncated sample is a weighted average of the individual estimates, with the weights given by the phenotypic variances. For simplicity, taking the unweighted average, gives a value of 0.025. Using Normal Theory on phenotypic truncation: the phenotypic variance is reduced by a fraction (1-k), where k is expected to be between 0.5 and 1 due to the high selection intensity (i.e. low proportion judged as 'IR' or 'R'); in contrast, the genetic variation is reduced by a fraction related to $(1\text{-}kh^2)$ where h^2 is the true unselected heritability. Therefore the theory suggests that the observed heritability *within the truncated fraction* is inflated by a fraction $(1\text{-}kh^2)/(1\text{-}k) > 1$. Hence the true heritability is likely to be less than that observed in the truncated sample and is predicted to be lower than 0.025. Such analysis strongly suggests that there is substantial environmental noise in response to an individual skin test, independent of infection. *Further opportunities to analyse the more recent VEBUS database which has relevant data and has grown substantially since the application would be welcomed.*

Therefore, drawing together the arguments, the observations from this study support the idea that the genetic variation lies primarily in susceptibility to TB and not in skin responsiveness independent of disease, and that any selection based on the testing outcomes e.g. u_D would have little impact on skin responsiveness independent of disease.

A further issue is whether or not the heritabilities are applicable to all regions where epidemiological factors may differ or where different spoligotypes of the M. *bovis* may be observed. There is some preliminary evidence that the spread of such spoligotypes is non-random (Smith et al. 2006). This is simply addressed, the estimates given here are expected to be representative of the observed national epidemic, averaging regional factors and spoligotypes in relation to their frequency in the national data, since parameters are estimated using data that is, in principle, drawn at random from the national epidemic. The analysis of regional data gave some interesting hints of perhaps a heterogeneity of extent of genetic variation, but no suggestion that sire effects are very different across regions. It has been a deliberate choice to focus this proof-of-principle study on first breakdowns as a cleaner approach, to avoid some of the possible immunological debates that might ensue otherwise. Therefore strictly, the data analysed are representative of breakdowns with no recent history (perhaps at the frontier of the epidemic) rather than all breakdowns. *This restriction could be usefully removed in further work.*

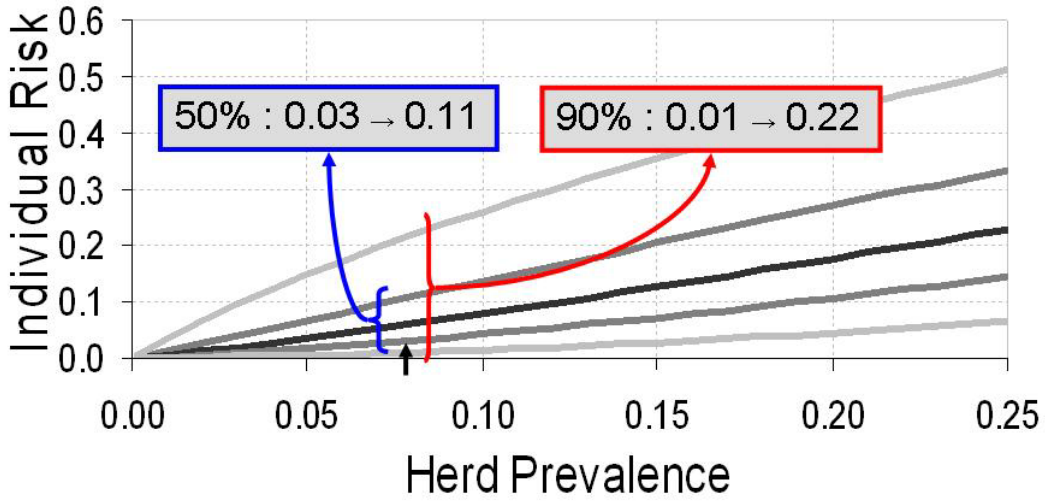
The approach conducted throughout this analysis has been to apply linear mixed models to threshold data. Throughout the development of the theory on such models there has been a concern on the extent of bias in the resulting estimates from the non-linearity of the link functions between the observed 0/1 scale and the liability scale. To protect against this, the key analyses on u_C and u_D presented here have been analysed using a variety of approaches: the results were analysed on the observed 0/1 scale and estimates scaled using Robertson & Lerner; sire models have been fitted rather than full pedigree, which have the benefit of fitting smaller variances on the liability scale and hence are expected to be less exposed to bias through non-linearity, although at a cost of loss of some of the genetic variation; simulation studies in which dummy data was generated with known parameters were analysed to assess the potential degree of bias. The conclusion from this extensive testing is that results are robust in this respect.

Finally, accepting there is genetic variation in TB susceptibility, what form might it take? Firstly the data is incapable of inferring whether the genetic variation includes complete resistance or only relative resistance. Complete resistance can only be inferred by a failure to become infected and this is difficult to establish. As stated previously the primary source of genetic variation is variation between sire half-sib families, and the well estimated breeding values are for commonly used sires. Suppose one such sire is completely resistant due to a dominant allele at some locus, then for this to be observed all offspring present in the analysis would need to inherit the dominant allele, an unlikely event in a large family. Alternatively, if due to a recessive allele then complete resistance in all the offspring would require all to inherit a similar allele from the dam, again an unlikely event unless this allele is very frequent. Further it is not possible to determine in what way the genetic variation is acting to reduce infection. For example the observations are compatible with a genetic difference in behaviour, but are also likely to be compatible with other barriers to infection. In the sections that follow the genetic variation is not assumed to confer complete resistance to TB only relative resistance.

Implications

The first implication is that within a herd there is heritable variation in individual risk for susceptibility to TB. An illustration of the extent of this is shown in Figure 2.

Figure 2. The 5, 25, 50, 75 and 95 per cent quantiles of the distribution of individual risk arising for a heritability on the liability scale of 0.18. The boxes highlight the fraction lying between 25 and 75 per cent quantiles (blue), and between the 5 and 95 per cent quantiles (red), for a herd prevalence of 7%. The value of 7% is chosen since this is close to the culling prevalence observed in the data summarised in Table 1. The values are derived from a Normal approximation to the models fitted.



Examining Figure 2, it shows that although for a typical herd there may be (say) 7% of cows culled during a breakdown due to TB, within the herd there are some cows much more likely to be culled due to their breeding value than others. The 5 lines show animals of increasing risk defined by how extreme they are in the population (defined by the quantiles). At 7% culling, 5% of the cows in the herd have a probability <0.01 of being culled, whilst 5% of the cows have a probability >0.22. The 'middle' 50% of the cows have probabilities ranging between 0.03 and 0.11. The 5 lines show how these probabilities change as the degree of culling increases or decreases.

Following from this, it is clear that genetics can play an important role in control strategies for TB in reducing the incidence of herd breakdowns and reducing the impact of herd breakdowns given their occurrence. It is not anticipated, nor even suggested (since it is assumed that the variation is not complete resistance) that it is the sole strategy. It is clear that with a disease such as TB a number of co-ordinated strategies need to be undertaken and genetics has the potential to make a substantial contribution as part of this wider effort. This is taken up in greater detail in the following section on 'Implementation and Supporting R&D'. Irrespective of the positive use of the genetic variation for control strategies, other strategies may benefit from examining their genetic implications, since as shown in Figure 2, individual risk can vary widely.

A second implication, of wider relevance than bovine TB, is that government surveillance data can be linked to industry databases to provide valuable evidence on population wide risk factors, including genetics. Furthermore this can be achieved relatively speedily, since these results were obtained within 12 months, from scratch. *One of the deliverables from this project will further speed up this process* in cattle, since cross-reference look-up tables have been set up which will facilitate the linking of key herd and animal identification fields across databases in any future study associated with TB or other diseases in cattle.

Implementation and Supporting R&D

The following provides a staged plan for implementation and R&D for exploiting the genetic variation in TB susceptibility.

Stage 1. Implementation with immediate effect. Immediately upon permission from Defra to discuss the results in detail, the existing EBVs (estimated breeding values) of sires can be used by breeding companies. The University of Edinburgh, SAC and its industry contacts (such as DairyCo) have considerable experience in delivering the appropriately measured messages necessary for delivering EBVs to farmers: for example, the team were behind the development and introduction of fertility indices and other initiatives in the dairy industry and, through Edinburgh Genetic Evaluations Service (EGENES), currently deliver all UK genetic evaluations for dairy, beef and sheep. There is sufficient information for breeding companies to match the promotion of bulls to regions with more or less TB. From conversations prior to this project we know that breeding companies such as Genus ABS and Cogent would participate in this, and we believe from dairy organisations that farmers would welcome steps they can take to reduce the risks of TB as well as introducing biosecurity measures.

The ***strengths*** of this Stage are: the implementation step is immediate; no further R&D is necessary; and it involves industry partners. The ***weaknesses*** of this Stage are: applicable only to dairy animals, not to beef; and the relevance of the EBVs calculated in this analysis will become dated quickly.

Stage 2. Consolidation and Networking. To address the weaknesses listed above it will be necessary to ensure the current evaluations carried out in this study are regularly updated and results integrated with other genetic evaluations. This would primarily involve the re-extraction of data using the established protocols and the running of the final models as a routine – hence the timescale of the project would be shorter than the current study to achieve this, but some recurrent funding would also be needed.

However additional benefits could be obtained from additional work to establish a more comprehensive database on which to base EBVs. This could be developed within UK by integrating the DARDNI database, and equally usefully by co-operation with the Republic of Ireland. The integration of evaluations would increase their accuracy and lead to faster progress, but some initial work would be required to check the genetic correlation between TB susceptibility in the Republic and the UK: indications are encouraging that this is likely to be high. Incorporating BCMS would in principle be useful by increasing the numbers of herd breakdowns for which 'Clear' animals can be identified, however whilst there is a provision for recording sire, it is not compulsory and *unless it is present the value of the records is limited*.

One benefit from networking with the Republic of Ireland would be a wider and stronger database for beef bulls than would be possible from the UK alone. In GB the beef herds are much smaller and the breeds themselves are less numerous. However there are some beef breeds that are more substantial in number than is typical for beef, such as the Limousin. Pedigree and performance records for these are stored on the BASCO database and an initial analysis could be done on these records to examine the opportunities. With small dairy breeds, parameters estimated from Holstein Friesian analyses are routinely used for EBVs after first checking that the evidence from within a small breed is consistent with the Holstein Friesian estimate, and this approach could usefully be applied to beef breeds.

An estimate of impact can be provided, at least for dairy. We shall assume that selection only takes place among males, and that the selection intensity for TB is 0.8, consistent with selecting bulls from the top half. Estimates of accuracy would be based on 8 to 10 daughters (derived from the approximate number of breakdowns recorded in dairy herds at present and the number of dairy bulls entering the population per year). Assuming a prevalence in the herds of 0.07 and a heritability on the liability scale of 0.20, the heritability on the 0/1 scale is ~0.06, as is σ_A, and the potential accuracy is ~0.33. The progress per generation is then ~ 0.5 x 0.8 x 0.33 x 0.06, or ~ 0.008, thereby reducing prevalence by about 10%. Progress would diminish as the epidemic prevalence diminishes. This is based on Normal approximations and a short simulation study would be worthwhile.

A further benefit is that it is likely that all aspects of this Stage could attract significant industry support.

The ***strengths*** of this Stage: integration of databases nationally; international initiatives; potential to strengthen involvement of industry bodies in the control strategy; more sustained progress in dairy herds; and, a preliminary step into beef herds. The ***weaknesses*** of this Stage: progress will diminish as the epidemic reduces; accuracy is only lowly moderate for dairy herds but is equivalent to other traits routinely evaluated and included in indices in dairy cattle breeding; and beef is patchily addressed.

Stage 3. Genome Wide Evaluation. A more sustained approach is to take advantage of dense SNP technology to derive genomic breeding values. A BBSRC project is already in progress in collaboration with AFBI in Northern Ireland in which a case control study is being undertaken, where cows with and without disease lesions will be genotyped using the Illumina 54k chip. The benefits of this approach are: (i) it is yet more firmly rooted in disease rather than skin tests, and (ii) it has the potential to derive a genetic predictor that is independent of the disease incidence, and and therefore the selection for TB resistance can continue even though the epidemic is under control. This may therefore offer a sustainable strategy. Whilst the greatest excitement of the technology exists with genome wide evaluation as described by Meuwissen et al (2000), the process of their development would also includes all QTL detection strategies based upon the SNP typing.

It would be valuable, to strengthen this study, to make it more definitive. This could be done by providing additional funds for genotyping and by help in sourcing suitable cases and controls. Depending on the genetic architecture, more work may be required in beef breeds.

The ***strengths*** of this Stage: a sustainable process of selection against TB susceptibility; an expected greater accuracy of selection leading to faster selection; directed towards disease and not to skin test outcomes; and uses cutting edge technology. The **weakness** of this Stage: the converse of cutting edge, genomic EBVs have considerable potential but the approach has yet to be explicitly demonstrated.

Summary of Further Research Proposed

Further research could then be proposed along the following lines:

1. Analysis of the VEBUS database to assess the skin responsiveness inheritance on unselected data, and further confidence building and testing of models from the existing data set. In the short timescale of the current project some of the points debated in the discussion above could be further clarified such as possible explanations of regional differences and including data on multiple breakdowns per herd, which could increase accuracy of evaluations and other traits, for example, somatic cell counts.

2. Assisting in the delivery of existing EBVs.

3. Preparing for genetic evaluations on a regular basis, including minor model development.

4. Extending linkage of databases and provision of genetic evaluations both nationally and internationally. This would include linkage and establishing genetic correlation with Republic of Ireland, and inclusion of BASCO for beef.

5. Participate in the BBSRC genome wide study by further genotyping and sourcing of cases and controls.

References to published material

9. This section should be used to record links (hypertext links where possible) or references to other published material generated by, or relating to this project.

Agresti, A. 2002. Categorical Data Analysis. John Wiley & Sons Inc. Hoboken, New Jersey.

de la Rua-Domenech, R., Goodchild, A., Vordermeier, H., Hewinson, R., Christiansen, K., Clifton-Hadley, R. 2006. Ante mortem diagnosis of tuberculosis in cattle: A review of the tuberculin tests, γ-interferon assay and other ancillary diagnostic techniques. *Research in Veterinary Science*, **81**: 190-210

Lesslie, I. W. and Hebert, C. N. 1975. Comparison of specificity of human and bovine tuberculin ppd for testing cattle. 3. National trial in Great-Britain. Vet. Rec. **96**: 338-341

Robertson, A., Lerner, I. M. 1949. The heritability of all-or-none traits - viability of poultry genetics. Genetics **34**: 395-411.

Smith, N.H., Dale, J., Inwald, J. et al. 2003. The population structure of Mycobacterium bovis in Great Britain: Clonal expansion. Proc. Nat. Acad. Sci. **100**: 15271-15275

Tallis, G.M. 1965. Plane-truncation in Normal populations. J. Roy. Stat. Soc. B **27**: 301-307

Woolliams, J. A. 2006. A note on the differential impact of wrong and missing sire information on reliability and gain. J. Dairy Sci. **89**: 4901-4902

ANNEXE 1. A glossary of technical terms used in the report.

Bivariate analysis provides for the joint analysis of two traits so that both correlations between breeding values and correlations between phenotypes for the two traits can be estimated. In many fields the methods used do not differentiate these correlations, yet they are commonly found to be different. A bivariate analysis is in contrast to a **univariate analysis** in which the traits are analysed without allowing for any correlation between them, or where assumptions are made that all possible correlations are equal to the phenotypic correlation.

Breeding value is the mean genetic value of an individual as a parent for a trait. It is estimated as twice the average superiority of the individual's progeny relative to all other progeny under conditions of random mating. The breeding value is the genetic value that is relevant to permanent population change, and constitutes the dominant part of the total genetic variance. In this study the term **genetic variance**, unless otherwise stated refers to the variance of the breeding values.

Cohort is a term used to describe a group of animals kept under similar conditions at the same time, a synonym for contemporary group.

Complementary log-log link function is one of three functions widely used to link the probability scale for a disease to the underying liability scale. These link functions are based on the Normal distribution (termed probit link), the logistic distribution (termed logistic link, and associated with log-odds), and the extreme value distribution (complementary log-log link). The latter arises naturally from some widely used models of survival times. However, empirically, in the types of models fitted here, with the prevalence observed here, the impact of the different link function is small.

Correlation coefficients express the strength of an assocation between two traits, as assessed by linear trends i.e. the degree to which one trait (y) is predictable from another (x) in the form y=bx+c. Values always lie between -1 and 1, a value of 0 indicates no predictability, the sign is the same sign as b and denotes whether the trend is increasing or decreasing. Values of magnitude <0.2 may be taken as small, >0.2 and <0.6 as medium and >0.6 as large.

Fixed effects are terms in a statistical model. They are most easily explained as not being random effects. See below.

Genetic correlation is the correlation between breeding values for two traits in a population. It is the genetic correlation that determines what direction and degree of change in selection programmes.

Genetic effects is used in the text as a synonym for breeding values.

Heritability is the fraction of phenotypic variance that is attributable to breeding values. It always lies between 0 and 1, where 1 indicates that all the variation observed can be explained by genetics and none by the environment, and 0 *vica versa*. The heritability will depend on the scale of measurement, for example some measurement might be much crude measures than others, in which case heritabilities will be lower. When a heritability is measured by direct analysis of 1's and 0's denoting diseased and disease-free individuals, then it is a common finding, and unsurprising biologically, for the **heritability on the 0/1 scale** to vary with the prevalence of the disease. However when the model of analysis considers a continuous underlying liability scale, and analysis estimates the **heritability on the liability scale**, it is found that the heritability is independent, or less dependent, on prevalence. Such analyses simultaneously use the 1's and 0's to estimate the liability of individuals and estimate what proportion of the variance in liability is attributable to genetics. Classification of heritability as **low**, **medium** or **high** is subjective and based on empirical observation so that low is perhaps <0.15, medium is >0.15 and <0.4 and high is >0.40.

Liability is a quantitative measure of how likely an individual is to be diseased. An individual's liability is considered here to be the sum of a genetic term and an environmental term, and if the sum of the terms exceeds a threshold then the individual is deemed to be diseased. The threshold is determined by the prevalance of the disease in the contemporary group.

Permanent environment is a term used to describe environmental influences that are persistent in differentiating the phenotype of one individual from another. These can be discerned when individuals are measured on more than one occasion. There are many possible origins of such variation. Consider, for example identical twins, one fed an energy rich diet and one fed a normal diet: one would expect to observe differences in, say, fatness between the twins that are persistent across repeated measurements, but these are not of genetic origin. The fraction of the total variance explained by the permanent environment is denoted in this study by $\mathbf{c}^2$. The value of c^2 for the variance due to permanent environment and so always lies between 0 and 1, and is analogous to the heritability for the genetic variance.

Phenotype is the observed value of a trait. It is a consequence of all the genetic and environmental influences and interactions affecting the trait, including errors in measurements. The **phenotypic variance** is a measure of how much variation is observed in the population.

Phenotypic correlation is the correlation between the phenotypes for two traits in a population.

Prevalence in a group describes the overall frequency of the disease in the group. The group may refer to a whole population or a contemporary group according to the context.

Quantiles help to describe the spread of a distribution. Say y is an 'x%' quantile of a distribution then the probability of a value sampled from the distribution being <y is x/100.

Random effects are terms that appear in a statistical model that are explicitly assumed to be drawn from a Normal distribution. To illustrate a consequence of this, assume a group of sires are fitted as a random effect, then the predicted effect of a sire is obtained by regressing back its progeny mean towards the overall mean by a degree that depends on the magnitude of the current estimate of sire variance and the error associated with the progeny mean. Thus sires with little information would have their progeny mean regressed back heavily towards the observed mean, whilst sires with lots of information have ther progeny mean regressed back much less. Therefore the predicted effect of a sire can only move away from the overall mean as information accumulates upon it and the predicted effect is unbiased.

ANNEXE 2. The record of the matching process used to identify ear tag numbers from the NMR database.

Record counts from the logfile created by the procedures are given below. The most recent is at the top and so the records must be read from the bottom up. Match processes 3 and 5 (c and e in the relevant section of Materials & Methods) gave the most matches in the HUK database and match process 9 (d in the relevant section of Materials & Methods) in the NMR database. The led to a total of 233616 matches across both databases.

comment	datedadded	rowsprocessed
Number of NMR recs inserted into TB_Sample_Data for found_code 13 =	11:37.2	3202
Number of NMR recs inserted into TB_Sample_Data for found_code 12 =	10:51.8	3634
Number of NMR recs inserted into TB_Sample_Data for found_code 11 =	07:46.8	0
Number of NMR recs inserted into TB_Sample_Data for found_code 10 =	52:39.5	2883
Number of NMR recs inserted into TB_Sample_Data for found_code 9 =	33:20.4	75199
Number of NMR recs inserted into TB_Sample_Data for found_code 8 =	23:05.8	0
Number of NMR recs inserted into TB_Sample_Data for found_code 7 =	10:15.3	0
Number of NMR recs inserted into TB_Sample_Data for found_code 6 =	59:43.6	1162
Number of NMR recs inserted into TB_Sample_Data for found_code 5 =	49:02.1	8608
Number of NMR recs inserted into TB_Sample_Data for found_code 4 =	40:15.6	21
Number of NMR recs inserted into TB_Sample_Data for found_code 3 =	29:20.9	55
Number of NMR recs inserted into TB_Sample_Data for found_code 2 =	13:51.3	1091
Number of NMR recs inserted into TB_Sample_Data for found_code 1 =	10:09.9	5793
Starting [usp_Find_TB_Animals_IN_NMR]	06:42.7	NULL
Number of HUK recs inserted into TB_Sample_Data for found_code 14 =	06:41.1	2413
Number of HUK recs inserted into TB_Sample_Data for found_code 12 =	04:01.7	2757
Number of HUK recs inserted into TB_Sample_Data for found_code 11 =	01:44.9	0
Number of HUK recs inserted into TB_Sample_Data for found_code 10 =	49:47.7	0
Number of HUK recs inserted into TB_Sample_Data for found_code 9 =	38:41.4	869
Number of HUK recs inserted into TB_Sample_Data for found_code 8 =	21:24.1	3
Number of HUK recs inserted into TB_Sample_Data for found_code 7 =	08:12.6	42
Number of HUK recs inserted into TB_Sample_Data for found_code 6 =	56:37.5	2647
Number of HUK recs inserted into TB_Sample_Data for found_code 5 =	46:08.0	56208
Number of HUK recs inserted into TB_Sample_Data for found_code 4 =	36:09.6	35
Number of HUK recs inserted into TB_Sample_Data for found_code 3 =	26:23.3	60466
Number of HUK recs inserted into TB_Sample_Data for found_code 2 =	17:24.8	1294
Number of HUK recs inserted into TB_Sample_Data for found_code 1 =	14:44.2	2980
Starting [usp_Find_TB_Animals_IN_HUK]	11:37.7	NULL

ANNEXE 3. Consideration of the nature of skin-test sensitivity and the extent of its genetic variation.

An outcome from considering the issue of 'skin-test sensitivity' is that this term could one or both of two different things.

- The 'skin-test sensitivity' may be *a property of an individual irrespective of disease*, the type that might be examined within the VEBUS database where field tests of individuals are recorded in disease free herds. A putative source of such a sensitivity may be an allergy to the skin-test in a differential way so as to interfere with test outcomes. Such a 'skin-test sensitivity' might be expected to affect both the epidemiological sensitivity (Se) and specificity (Sp), so that an individual whose immune system reacts in this way might be more likely to be seen as as a reactor.
- Alternatively the 'skin-test sensitivity' may be regarded as the *probability the test will be positive <u>given</u> that the disease is present*, so that it is seen as an interaction of the individual with the pathogen, and not as a property of an individual.

It is the first of these that received most thought in the analyses and although other more general arguments were made they were more vague. There is a critical difference between the two of these: in the first all animals are relevant since all animals may give information about innate reactions; whereas in the second only those with disease provide information that is relevant to the genetics of this type of response.

The hypothesis of concern in the second case is that what is identified as a heritability in the susceptability to TB, based upon the comparison of confirmed cases and the remainder, could in fact reflecting the heritability of the probability of response. This requires two things to be credible:

- Se<1 for their to be a contrast between identified cases and non-identified cases.
- The pattern of inheritance to be strong enough to be exhibited when cows that are free of disease will mask the expression of the trait in the negative test group. The fraction of cows free of the disease is 1-p, where p is the prevalence.

Re-casting the problem in this way, shows the mathematics to be analogous to exposure. For the issue of exposure, a fraction (1-e') are not exposed to the disease and, of the fraction e' that are exposed, a fraction p' are seen to become diseased. For 'skin-test' sensitivity, a fraction (1-p) are not expressing the trait and, of the fraction p that are, a fraction Se are seen to become diseased. Using this analogy the work of Bishop and Woolliams (manuscript in preparation) can be used to assess the observed heritability of 0.19 found in the study.

So consider the trait of *probability the skin-test will be positive <u>given</u> that TB is present*. The observed heritability when using confirmed cases will depend on prevalence p and sensitivity Se, but not Sp since Sp=1. The observed prevalence will be pSe. If the observed prevalence is constrained to 0.07, then true prevalence p is 0.07/Se. Table 1 shows the value of h^2 expected for different values of Se for the entire skin-test procedure over a breakdown when the trait is either <u>completely</u> <u>heritable</u>, i.e. $h^2 = 1$, or $h^2=0.5$

Table 1. The heritability that is predicted to be observed for *probability the skin-test will be positive <u>given</u> that TB is present* from conducting an analysis of confirmed cases and the remainder, when assuming the true heritability is either 1 (so completely heritable) or 0.5.

Sensitivity, Se	**0.6**	**0.7**	**0.8**	**0.9**
True $h^2 = 1$				
Observed h^2	0.113	0.067	0.033	0.010
True $h^2 = 0.5$				
Observed h^2	0.056	0.033	0.017	0.005

The implication of Table 1 is that the values shown are highly unlikely to be consistent with the observed value of 0.19 and its associated standard error (0.04). The range of values for heritability may be considered to be highly optimistic, and the Se ranges down to values that are pessimistic for the complete process, so the values presented are conservative for making the inference. Therefore we might infer that the sensitivity to the skin-test in this form is not responsible for the large majority of the genetic variation observed in u_D. If such variation exists then it is expected to constitute be a small fraction of the variation and to have only small genetic correlation with u_D. Therefore population changes would expect to decrease the prevalence of TB, and if the number of cases decreases it might be expected to increase the fraction of cases that are silent, but with no substantial increase in number.