

TB in cattle

Defra's response to the independent audit on the quality and completeness of data collection in the Randomised Badger Culling Trial.

CONTENTS

SECTION

1. EXECUTIVE SUMMARY	3
2. BACKGROUND	4
3. APPOINTMENT OF AUDITOR	4
4. COMMENTS ON AUDIT REPORT	4
5. RECOMMENDATIONS IN REPORT (MAIN BODY)	5
6. RECOMMENDATIONS IN REPORT (ADDITIONAL) & ISG RESPONSE	6
7. REFERENCES	7
ANNEX 1	9

SECTION 1: EXECUTIVE SUMMARY

1.1 The fieldwork and laboratory testing of the Randomised Badger Culling Trial (RBCT) ended in March 2006. Results from the RBCT are being assessed by the Independent Scientific Group on Cattle TB (ISG) and they expect to publish their final report and recommendations to Ministers in June 2007. Findings¹ up to 2005 from the Proactive treatment of the RBCT have been published and a detailed account² of the Reactive treatment, which ended in 2003, is also available.

1.2 The RBCT started in 1998 and has been an important part of the Government's research programme to investigate the control of tuberculosis (TB) in cattle. The design of the RBCT and its implementation have been described in the reports³⁻⁷ of the Independent Scientific Group on Cattle TB (ISG) and on the Internet site⁸ of Defra. A number of independent audits have focussed on key elements of the RBCT; these have been published together with a Departmental response, where applicable. The audits have covered the effectiveness of the surveying for badger setts⁹, the statistical validity of the trial¹⁰⁻¹², badger post mortem examination¹³, bacteriological culture procedures¹⁴ and the humaneness of badger dispatch procedures¹⁵⁻¹⁹.

1.3 After a proposal by the ISG, the Department decided to commission an independent audit, to monitor RBCT data quality and completeness, which took place in 2005-6 and is the subject of this report.

1.4 Dr Martine Wahl undertook the audit and her report is at Annex 1.

1.5 The Department welcomes the Auditor's findings (see Annex 1, Executive Summary) that the RBCT documentation was of a high standard, data management had been effective and that data extraction had been performed correctly. Additionally, the Department welcomes the finding that when potential difficulties had been identified by the Auditor, Departmental staff had investigated and that, subsequently, the Auditor was able to state that these matters had not affected the primary conclusions of the Trial.

1.6 In the main body of the Auditor's final report there were six recommendations. Five of the recommendations (1-2 & 4-6), on data quality and archiving matters, the Department accepts and has either made the changes required or will shortly complete these. However, the third recommendation, a proposed extension of the database to include all occupiers within Trial Areas, cannot practically be implemented and is therefore rejected. A detailed response is in Section 5.

1.7 The Auditor added (page 24, after paragraph 7.6) to her final report a Section "Further recommendations not previously made in the report" and, whilst the Department was engaged in making some of the changes recommended (in the main body of the report), the Auditor was able to further assess and make additional comment. Responses on this can be found in Sections 5 and 6.

SECTION 2: BACKGROUND

2.1 Ministers announced, on 5 February 1998, that there would be a scientific study (the RBCT) to investigate the link between TB in cattle and badgers. The study started in August 1998 and followed from recommendations of two independent committees, the first headed by Professor Sir John Krebs and the second by Professor John Bourne. The description of the RBCT has been made widely available³⁻⁸, and certain statistical analyses have already been published¹⁻².

2.2 The Department realised at the outset of the RBCT that it would be extremely important to ensure that all aspects of the Trial met the highest standards of scientific probity. To assure this, when the range of independent audits⁹⁻¹⁹ was commissioned, an undertaking was given that audit reports and any Departmental response to these would be published in full. The audit report at Annex 1 and this document meet that undertaking.

SECTION 3: APPOINTMENT OF AUDITOR

3.1 Dr Martine Wahl MD was appointed in October 2004 after an open competition.

3.2 Dr Wahl, based in Basel Switzerland, is an experienced Auditor of clinical research and has previous experience of audit in the bovine tuberculosis field from her appointment by the Department in 2003 as Auditor of the TB99 epidemiological surveys.

SECTION 4: COMMENTS ON AUDIT REPORT

4.1 This Section of the Department's response sets out general comments on the audit report. Paragraph/Section numbers are included for ease of reference.

Methods.

4.2 The Department accepts the audit methods (Section 4) used by Dr Wahl and is grateful for the orderly, meticulous and most helpful approach taken.

Departmental staff.

4.3 The Department is pleased that the Auditor was able to comment (paragraph 6.1.1) that Defra Wildlife Unit (WLU) and Veterinary Laboratories Agency (VLA) staff were well motivated, had a good understanding of their duties, functioned well as a team and were helpful during the audit process. The Department notes no adverse comments have been made concerning any staff member.

ISG.

4.4 During the audit, the Auditor met members of the ISG (*vide infra*, paragraph 6.2) responsible for RBCT design and data analysis and asked (paragraph 5.1.3) for clarification concerning the independence of the ISG with regard to its involvement in the statistical analysis. The Department is pleased that the Auditor was able to comment that she believed the ISG had no direct involvement in the collection of the data. The Department agrees with the Auditor that the planning, implementation and interpretation of the statistical analysis was and is appropriately independent.

SECTION 5: REPORT RECOMMENDATIONS (MAIN BODY OF REPORT)

Recommendation 1 "In order to aid future analysis by individuals who have not been involved with the trial data, I recommend that a project and protocol document that contains details and definitions of key data and an explanation of how they can be recreated from the raw data is drawn up: e.g. the definition of a trial breakdown. This document should also include a summary of the changes over time. Such a document will form an invaluable aid for future analysis". (Paragraphs 6.1.2.1, 6.1.3.2, 6.2.2).

5.1 The Department agrees with this recommendation. A document containing details of key data and how these may be recreated from the raw data, along with the addition of a summary of changes over time, will be drawn up.

Recommendation 2: "All data entry forms should be examined to see if further cross-consistency checks can be identified. If such checks are identified appropriate queries should be run against the data, allowing the data to be cleaned. It is essential that the data is as clean as possible for future usage, particularly at the point it is opened up to 'outside' analysis." (Paragraph 6.2.2).

5.2 The Department agreed with this recommendation. This work has now been completed and checked by the Auditor.

Recommendation 3: "There should be a final process of checking the GIS database against the trial database for cross-consistency in terms of trial occupiers. Furthermore the GIS database should be extended to include all occupiers within trial areas, whether or not they had been signed up to the trial. This process would provide validation of the trial population figures". (Paragraph 6.3.1.2).

5.3 The Department has concluded that the absence of data on boundaries for occupiers of unsigned land (a proportion of which is non-agricultural) within Trial Areas means that such Areas cannot realistically be broken down further than these already are. A proposed extension of the database, to include all occupiers within Trial Areas, cannot be practically or economically

implemented and this recommendation is therefore rejected. The Department believes the GIS database is as complete as is necessary for present and any future analysis.

Recommendation 4: "As data on changes to farm details, such as dates when farms were archived, exists on VetNet, an attempt should be made to recreate occupiers that were deleted from the trial database before the process of archiving occupiers was implemented." (Paragraph 6.3.1.3).

5.4 The Department agreed with this recommendation. The Auditor comments that this work has now been completed.

Recommendation 5: "There should be validation of the 'within trial area breakdown list' against data held locally at the local AHDOs. When a validated list of breakdown list is compiled it should form part of the trial database, rather than having to be recreated for each new analysis. The breakdown list should be updated at regular intervals." (Paragraph 6.3.2.3).

5.5 The Department agreed with this recommendation. Through the offices of the Animal Health (formerly, the State Veterinary Service) and VLA, this has been done to the extent that it is possible to do.

Recommendation 6: "For all badgers that formed part of the 'extended culture' work there should be a check that the culture result that exists on the trial database, is the culture result according to the standard protocol. This is because of concerns raised at the WLUs that some of the culture results forwarded to them for data entry may in fact have been the extended culture results." (Paragraph 6.3.3).

5.6 This work has now been completed and checked by the Auditor.

SECTION 6: REPORT RECOMMENDATIONS (ADDITIONAL) & ISG RESPONSE

6.1 The Auditor completed her report by adding a Section (page 24) entitled "*Further recommendations not previously made in the report*". The Department now responds to these four additional recommendations, but regarding the first of these, namely

"In future, audit activities should best be initiated at the planning stages of a trial rather than when a trial is already well in progress.",

the Department does so by endorsing the response from the ISG (reproduced below), both to this recommendation and the audit in general.

ISG RESPONSE

6.2 "The ISG welcomes the report and has been reassured by it, in particular by the reassurance that the Trial conclusions are in no way compromised by

any of the issues raised. The ISG has made the point that, from the outset, it has emphasised the importance of external audit and quality management of aspects of its work, and highlighted this (paragraphs 14.1 - 14.4) in their first report³ to MAFF in 1998. With reference to this audit report, therefore, the ISG was particularly interested to note the recommendation that "audit activities should best be initiated at the planning stages of a trial, rather than when a trial is already in progress". The early establishment of Trial audit procedures was advocated by the ISG in 1998."

6.3 Regarding the other three recommendations, the Department makes the following comments.

"At some stage Defra will have a responsibility to "sign off" the data, after which no changes to the data can be made."

6.4 The Department agrees with this recommendation. The database will be signed off formally once the ISG has completed its programme of planned analyses and before it disbands in the Summer of 2007.

"Defra has also a responsibility to ensure that the data, including source data, are archived in such a way that they can be retrieved at short notice."

6.5 The Department agrees with this recommendation and has appropriate work in hand to ensure that all data and materials from the RBCT, which it sees as valuable resources, are archived and made available. The Department will encourage bona fide researchers to make use of these in order to obtain the maximum benefit from its investment in carrying out the RBCT.

"Defra needs to maintain control over the future release of the data. It should consider forming a review panel, that includes those who have been responsible for the data, to look at any communications prior to publication, that result from the release of this data."

6.6 The Department agrees with this recommendation and believes it should take an interest in any work using RBCT data, wishing to be kept apprised of results as these emerge. The Department is currently deciding how it can best ensure effective monitoring of, and input to, results that will come out of work using RBCT data and/or materials.

SECTION 7: REFERENCES

1. Donnelly, C.A., Woodroffe, R., Cox, D.R., Bourne, F.J., Cheeseman, C.L., Clifton-Hadley, R.S., Gao Wei, Gettinby, G., Gilks, P., Jenkins, H., Johnston, W.T., Le Fevre, A.M., McInerney, J.P. and Morrison, W.I. Positive and negative effects of widespread badger culling on cattle tuberculosis. *Nature*, 2006, 439, 843-6.

2. Le Fevre, A.M., Donnelly, C.A., Cox, D.R., Bourne, F.J., Clifton-Hadley, R.S., Gettinby, G., Johnston, W.T., McInerney, J.P., Morrison, W.I. and Woodroffe, R. The impact of localised badger culling versus no culling on TB incidence in British cattle: a randomised trial” <http://www.defra.gov.uk/animalh/tb/isg/pdf/lefevre1005.pdf>.
3. Towards a Sustainable Policy to Control TB in Cattle. First Report of the Independent Scientific Group on Cattle TB. MAFF Publications, PB3881, 1998 (London).
4. An Epidemiological Investigation into Bovine Tuberculosis. Second Report of the Independent Scientific Group on Cattle TB. MAFF Publications, PB4870, 1999 (London).
5. An Epidemiological Investigation into Bovine Tuberculosis. Third Report of the Independent Scientific Group on Cattle TB. Defra Publications, PB 5801, 2001 (London).
6. An Epidemiological Investigation into Bovine Tuberculosis: towards a Science-based Control Strategy. Fourth Report of the Independent Scientific Group on Cattle TB. Defra Publications, PB 10138, 2005 (London).
7. Report of the work of the ISG and its Published Findings in 2005. <http://www.defra.gov.uk/animalh/tb/isg/5th-isgreport.pdf>
8. <http://www.defra.gov.uk/animalh/tb/culling/index.htm>
9. Cresswell, W. Report of the Independent Audit of Surveying and Badger Social Group Territory Delineation Procedures in the Randomised Badger Culling Trial and MAFF’s response. MAFF Publications, PB5497, 2001 (London).
10. Mollison, D. First Report of the Statistical Auditor on the Badger Culling Trial. MAFF Publications, PB5385, 2000 (London).
11. Mollison, D. Report of the Statistical Auditor on the Badger Culling Trial. 2004; <http://www.defra.gov.uk/animalh/tb/culling/p5aud.htm>
12. Mollison, D. Report of the Statistical Auditor on the Badger Culling Trial. 2005; <http://www.defra.gov.uk/animalh/tb/culling/p5aud.htm>
13. Hall, G. Report of the Independent Auditor on the badger post mortem procedures used in the Randomised Badger Culling Trial and Defra’s response. Defra Publications. PB9702, 2004 (London).
14. Corbel, M. Report of the independent audit of bacteriological culture carried out on tissue samples collected during the randomised badger culling trial and Defra’s response. PB10204, Defra, 2004 (London).

15. Kirkwood, J.K. Report of the Independent Auditor on the Humaneness of the Despatch Procedures used in the Randomised Badger Culling Trial and MAFF's Response. MAFF Publications, PB5325, 2000 (London).

16. Ewbank, R. Report of the second independent auditor on the humaneness of dispatch procedures used in the randomised badger culling trial and Defra's response. Defra Publications, PB8253, 2003 (London).

17. Ewbank, R. Report of the third independent audit on the humaneness of dispatch procedures used in the randomised badger culling trial and Defra's response. PB9957, Defra 2004 (London).

18. Anderson, J.C. The fourth independent audit on the humaneness of dispatch procedures used in the randomised badger culling trial and Defra's response. PB11329, 2004, Defra 2005 (London).

19. Anderson, J.C. The fifth independent audit on the humaneness of dispatch procedures used in the randomised badger culling trial and Defra's response. PB11908, Defra 2006 (London).

**Department for Environment, Food and Rural Affairs
April 2007**

ANNEX 1 (next page)

Text of Audit Report.



Appendix 1

An audit of the Randomised Badger Culling Trial (RBCT)

Prepared by: Martine Wahl MD
Principal Auditor/CEO
Clinical Research & Communication
Maiengasse 2
4056 Basel Switzerland

Key Words: Planning, communication, coordination, consistency, documentation, team approach

Date: November 2006

EXECUTIVE SUMMARY	4
1. INTRODUCTION	5
2. ROLES AND RESPONSIBILITIES.....	5
2.1 ISG.....	5
2.2 DEFRA ISG SECRETARIAT	5
2.3 DEFRA WILDLIFE UNITS (WLU)	6
2.4 DEFRA STATE VETERINARY SERVICE (SVS)	6
2.5 VETERINARY LABORATORIES AGENCY (VLA)	6
3. RATIONALE.....	6
4. MATERIAL AND METHODS	6
4.1 MATERIAL	6
4.1.1 DEFRA provided	6
4.1.2 VLA provided.....	7
4.1.3 WLU provided.....	7
4.1.4 ISG provided.....	7
4.2 METHOD	7
4.2.1 Data on badgers culled	7
4.2.2 Data on where the RBCT has taken place	7
4.2.3 Data on badger Setts and social groups	7
4.2.4 Data on traps set and trapping operations	7
4.3 AUDIT PLAN.....	7
4.3.1 Quality	8
4.3.2 Data management	8
4.3.3 Key data	8
4.3.3.1 Population data and recruitment	8
4.3.3.2 Herd breakdown	8
4.3.3.3 Badger TB prevalence	8
4.3.4 Extraction of data for analysis	9
5. VISITS MADE	9
5.1 VISITS TO WLU.....	9
5.1.1 Visits to WLU.....	9
5.1.1.1 Polwhele.....	9
5.1.1.2 Aston Down.....	9
5.1.2 Visits to the VLA (Weybridge)	10
5.1.3 Meeting with the ISG.....	10
6. FINDINGS	10
6.1 QUALITY	10
6.1.1 Staff	10
6.1.2 Planning and trial design	10
6.1.2.1 Protocol	11
6.1.2.2 Database	11
6.1.2.3 The Badger identification system	11
6.1.3 Documentation	12
6.1.3.1 General comments	12
6.1.3.2 Protocol amendments	12
6.1.4 Communication.....	13

6.2	DATA MANAGEMENT	13
6.2.1	Data management activities	13
6.2.2	Data cleaning	14
6.3	KEY DATA	15
6.3.1	Population data	15
6.3.1.1	Recruitment	15
6.3.1.2	Identification and recruitment of occupiers.....	15
6.3.1.2	Trial GIS Data.....	17
6.3.1.3	Changes in ownership or withdrawal from the trial	18
6.3.2	Herd Breakdown data	19
6.3.2.1	Collection of breakdown data.....	19
6.3.2.2	RCT 17 Data	20
6.3.2.3	AHDO data on breakdowns in trial areas.....	20
6.3.3	Badger TB prevalence data	21
6.4	ISG ANALYSIS	21
6.4.1	Extraction of trial database data	22
7.	<u>AUDITOR'S RECOMMENDATIONS</u>	23
7.1	RECOMMENDATION 1: PROJECT AND PROTOCOL DOCUMENT (SEE 6.1.2.1, 6.1.3.2, 6.2.2)	23
7.2	RECOMMENDATION 2: DATABASE "CLEANING" (SEE 6.2.2)	23
7.3	RECOMMENDATION 3: GIS DATABASE (SEE 6.3.1.2)	23
7.4	RECOMMENDATION 4: CHANGES IN OWNERSHIP OR WITHDRAWAL (SEE 6.3.1.3)	23
7.5	RECOMMENDATION 5: AHDO DATA ON BREAKDOWNS IN TRIAL AREA (SEE 6.3.2.3)	23
7.6	RECOMMENDATION 6: BADGER TB PREVALENCE DATA (SEE 6.3.3)	24

Executive Summary

The audit of the Randomised Badger Culling Trial (RBCT) data was proposed by the ISG and conducted at the request of DEFRA. Because of the extent and volume of data the auditor focused her attention on those aspects of the data quality and completeness that have been, and will be, important in the conclusions to be drawn from the trial.

All staff interviewed in the process of the audit were motivated and efficient.

Documentation relating to the RBCT was of a high standard. Further analysis by people not familiar with the RBCT would benefit from having a comprehensive project and protocol document.

Data management has been effective. Further identification of cross-consistency checks were highlighted as an area that would improve the reliability of the final version of the database. These checks have now been carried out and any data issues raised have been addressed and dealt with to the auditor's satisfaction.

Trial breakdowns were obtained by linking the RBCT data to VetNet data and not by recording herd breakdown as part of the trial database, this has led to a small number of breakdowns being misclassified as RBCT breakdowns. However these numbers are so small that they have had no effect on the trial conclusions. The drawing up and inclusion in the RBCT database of a validated list of breakdowns forms one of the auditor's recommendations.

The auditor is satisfied that the data extraction phase of the analysis was performed correctly.

Wherever potential difficulties have been identified the auditor is satisfied that these have been explored during the analysis phase, and have not affected the primary conclusions of the trial.

1. Introduction

The Krebs Report in 1997¹, which was a comprehensive review of bovine TB in cattle and badgers, highlighted the need for a large controlled experiment to establish the relationship between bovine and badger TB. This project, the Randomised Badger Culling Trial (RBCT), was initiated in 1998 under the guidance of the Independent Scientific Group on Cattle TB (ISG).

The RBCT was designed by the ISG and initiated by DEFRA to study the effect of three different treatments (badger culling policies) on herd breakdown rates. The three treatments were: the proactive culling of badgers in a given trial area; the reactive culling of badgers around infected farms, and no culling (a control treatment known as “survey only”). In total there were 10 groups, each group consisting of 3 treatment areas of approximately 100 square km in area. Each group of three treatment areas was known commonly as a ‘triplet’. The first of these 10 triplets started operation in December 1998 and the final triplet came into operation in December 2002. A number of factors, including the foot and mouth disease epidemic, delayed the original timetable of the RBCT.

In November 2003 an interim analysis of the RBCT showed that the reactive treatment strategy appeared counterproductive and led to an increased risk of TB in cattle in reactive areas and could not be considered as a viable policy option. The reactive treatment ceased as a treatment from this time on, and ever since, reactive areas have become, in effect, control areas.

The ISG, appointed to advise DEFRA, made it clear that it wished to see all aspects of the RBCT audited. The audit of the RBCT trial data forms part of the wider audit of the whole of the RBCT.

2. Roles and responsibilities

2.1 ISG

The ISG is an independent group set up by the government to advise on the design and running of the RBCT. The ISG has been (and will be) responsible for the analysis of RBCT data.

2.2 DEFRA ISG SECRETARIAT

The ISG secretariat is comprised of DEFRA officials who give administrative support to the ISG. In particular they are responsible for the administrative support to the ISG including the organization of the data handling meetings and taking the minutes at these meetings to ensure discussion is recorded and actions identified.

2.3 DEFRA WILDLIFE UNITS (WLU)

The WLU is responsible for the day to day running of the trial in the field as well as RBCT data collection and the collation of the trial database.

2.4 DEFRA STATE VETERINARY SERVICE (SVS)

The SVS is responsible for TB control in cattle in Great Britain, including within trial areas, and also for the VetNet database. It is now a DEFRA agency.

2.5 VETERINARY LABORATORIES AGENCY (VLA)

The VLA is responsible for the design, writing and maintenance of the trial database.

As part of this audit, members of each organization (with the exception of the SVS) were interviewed.

3. Rationale

The ISG first report, July 1998, paragraph 14.4, recommended that an external audit of MAFF (DEFRA) data should be done at the beginning of the trial².

The rationale for the data audit is to establish reasonable assurance that the conclusions or recommendations made at the end of the trial are based on the data collected which is as complete, accurate and reliable as possible, and that all steps involved in the extraction of the data for the analysis are fully documented and hence reproducible.

The detailed statistical analysis has been audited separately.³

4. Material and Methods

4.1 Material

All pertinent material was available for the auditor.

4.1.1 DEFRA provided

- Towards a sustainable policy to control TB in cattle (July 1998)².
- The administrative Standard Operating Procedures of the WLU (27 July 2004, 17 May 2005)
- Independent Scientific Review of the RBCT and associated epidemiological research (4th March 2004)⁴.

4.1.2 VLA provided

- The RBCT Access database (June 22nd 2004), Routine VLA data checks and all printouts required by the auditor.

4.1.3 WLU provided

- Standard Operating Procedures: Surveying (version 2.2)
 - Cage trapping of badgers (version 2.3)
 - Post Mortem examination of badgers (version 3.1),
- all the above were ratified on 7 April 2004

4.1.4 ISG provided

- Numbers of confirmed breakdowns and baseline herds by trial area (proactive and “survey-only” areas).

4.2 Method

A very comprehensive data set has been accumulated. A summary of the data collected during the 7 years of the trial is given below:

4.2.1 Data on badgers culled

This includes basic epidemiological information such as weight, age, sex together with details of their *post mortem* (PM) examination and culture status (positive or negative for *M. bovis*).

4.2.2 Data on where the RBCT has taken place

This includes data collected on occupiers (landowners) and also Geographical Information System (GIS) data on land areas.

4.2.3 Data on badger Setts and social groups

4.2.4 Data on traps set and trapping operations

4.3 Audit Plan

Because of the volume and the complexity of the data collected over the last 7 years, it would be impossible to audit all aspects of the RBCT data. The approach taken here has been to focus on the following aspects of the RBCT: quality, data management and key data.

4.3.1 Quality

Quality was assessed by checking the steps that have been taken to ensure reliable, complete and accurate data. It covers such aspects as: the expertise of the staff in place, planning and trial design, documentation and communication.

4.3.2 Data management

Data management concerns the steps that were taken to ensure that data were complete, consistent, clean and ready for analysis.

4.3.3 Key data

Key data are those essential to the analysis and its interpretation. Three key data areas were identified.

4.3.3.1 Population data and recruitment

For the trial, the population data is the number of occupiers* within trial areas. These occupiers may or may not be recruited into the trial, depending on the owner's acceptance or refusal. The analysis of cattle incidence data has been performed regardless of recruitment status i.e. it makes no distinction between occupiers that were recruited to the trial and those that were not.

The original population data was obtained from VetNet which was then 'downloaded' into the trial database (with additions manually added where new occupiers were recruited). The trial GIS database also contains details of the population data.

** An occupier is the name given to a land parcel or group of land parcels in the ownership of one business or individual, normally referred to as a farm. – The majority of occupiers are farms. There may be several parcels for one occupier in many instances.*

4.3.3.2 Herd breakdown

The outcome measure is the number of herd breakdowns in trial areas. Because these data were recorded not as part of the trial database, but on VetNet, there needs to be assurance that the way in which they have been calculated is reliable and accurate. One way of assessing the quality of these data is to compare them with source data held at the SVS Animal Health Divisional Offices (AHDOs).

4.3.3.3 Badger TB prevalence

Badger TB prevalence is likely to be crucial to future analysis. To date it has been used to determine the spatial association of TB in badgers and cattle⁵ and to describe temporal trends⁶.

4.3.4 Extraction of data for analysis

Because the data set, on which analysis have been performed, involves extraction and linking of the data to other data sets, there needs to be confidence that this has been done using the correct method; and in such a way that it is repeatable.

5. Visits made

The two Wildlife Units (WLU), Aston Down and Polwhele were visited, Polwhele once and Aston Down on two occasions. The VLA was visited on three occasions. A meeting was held with the ISG members having responsibility for the trial analysis.

5.1 Visits to WLU

Essentially these visits were to comprehend the process of data collection, to interview the staff, and to assess the consistency of approach between the two WLU's. At these visits the auditor compared data entered in the database with source data.

5.1.1 Visits to WLU

5.1.1.1 Polwhele

During the visit, the WLU team explained the data flow processes, from trapping to culture results, to the auditor. The organization of the team was clearly described by the Executive Officer (EO) responsible for trial data.

A folder, containing the local desk instructions was handed over to the auditor. These desk instructions were used in conjunction with the administrative procedures document version 2.1 dated 17 May 2005. They are clear and concise. The EO responsible for trial data took the auditor through the data associated with a number of the data capture forms - RCT4 (Sett survey), RCT 6 (Trapping), RCT 7 (badger carcass label). The team showed the auditor records of the quality control processes that were followed, and took her through one example.

The auditor also visited VLA Polwhele, responsible for PM examination and preparation of culture samples.

5.1.1.2 Aston Down

In Aston Down the auditor went through the process of data entry and quality control with the WLU trial data manager. She also discussed the interaction between the WLU and the VLA. The auditor followed the process of entering data on to the database with the administrative staff. The trial data manager described, in detail, the function of the unit, and the senior quality advisor described his role and the quality measures, from field collection of data to data entry, that were in place.

5.1.2 Visits to the VLA (Weybridge)

At the VLA the trial database, and the details of the data it contains, were described.

Data handling procedures were described very thoroughly by the VLA data manager in charge of the database.

A document describing the RBCT Access data base version 1.3, dated 18 August 2004, was given to the auditor. This document clearly describes all the steps for data entry, and was written by the VLA data manager. In addition, update changes have been documented and can be easily retrieved.

5.1.3 Meeting with the ISG

A meeting with the ISG members responsible for the trial analysis, was held towards the end of the audit. In addition to the matters covered in this meeting mentioned in 6.4., and in the conclusions, the auditor asked for some clarifications concerning the independence of the ISG with regard to its involvement in the RBCT statistical analysis. The ISG reply was that "they were to perform the primary analyses (that is, the analysis of cattle TB incidence data) because they were independent of Defra. Furthermore, the results were to be kept confidential until such time as they were sufficiently informative to be released to DEFRA ministers".

Auditor's comments:

The auditor was satisfied with the answers she received. She feels that the group carrying out the statistical analysis has had no direct involvement in the collection of the data. This ensured that the planning, implementation and interpretation of the statistical analysis was appropriately independent.

6. Findings

6.1 Quality

6.1.1 Staff

The staff was well trained and followed the administrative procedures. It was clear from each of the visits made (to the WLUs and the VLA) that all the staff, in each of the centres, were well motivated, had a good understanding of their duties, and functioned well as a team. Each team was very efficient and helpful in retrieving and checking data.

6.1.2 Planning and trial design

The critical stage for any trial is the planning phase. It is one of the most important periods for the success of a trial. In the experience of the auditor the success of a

study is greatly associated with proper planning, including a good protocol; good data capture record forms and a database developed together with the various team members.

6.1.2.1 Protocol

The ISG provided in its first report to DEFRA in July 1998², its recommendations as to how they would like DEFRA to run the RBCT. A large number of operational procedures were prepared by the ISG and DEFRA for field operations and included a plan for statistical analysis.

Auditor's comment:

The auditor was surprised that DEFRA had not collated this information into a project and protocol document

Auditor's recommendation:

The auditor recommends that a project and protocol document is created. (See recommendation 1)

6.1.2.2 Database

The RBCT database was initially based on the previous badger database used for badger removal operations (BROs). This was not ideal and it would have been preferable had the database been written from the onset to conform more directly with the requirements of the RBCT. The database was re-written in a new programming language in 2003 and at this time new validation was added. However the structure of the database remained largely the same. The staff at the WLU were very complimentary of the changes made in 2003, since the new system was much more "user friendly".

6.1.2.3 The Badger identification system

Whilst examining badger data the auditor was struck by two facts. First, the format of the badger ID changed during the life of the trial from being a five-figure digit (e.g. 12345 to two characters and 7 digits (e.g. AD 1234567) and secondly, the fact that some badger identifiers did not appear on the database.

Auditor's comment:

On the first point, it would have been ideal for the identifier to have stayed the same throughout the life of the trial, and this change may present problems to future data analysts;

Auditor's recommendation:

Details of the change should appear in the summary of changes over time document which would be part of the project and protocol document. (see 6.1.3 and recommendation 1).

On the issue of missing identifiers, table 1 shows a typical example with the identifiers AD07453 and AD07460 missing. This raises the issue of whether the missing identifiers represent missing badgers i.e. badgers that were captured but their data did not get onto the database or unallocated labels. The number of missing identifiers is not small. The auditor was assured that the reason for the missing identifiers was that identification tags were allocated before the badger trapping operation and the missing tags represent those that were not used.

Auditor's comment:

A better system would have been one that accounted for all badger carcass labels, and tags not used should have been returned, labeled as not used and accounted for.

Table 1 Example of non-sequential badger identifiers

Badger Id	VLA Lab	BdTaken
ad07450	17	24/07/2004
ad07451	21	09/08/2003
ad07452	21	09/08/2003
ad07454	27	12/08/2003
AD07455	27	09/09/2003
AD07456	21	10/08/2003
ad07457	26	07/09/2004
AD07458	26	15/08/2003
ad07459	21	10/08/2003
AD07461	21	09/08/2003
AD07462	12	14/08/2003

6.1.3 Documentation

6.1.3.1 General comments

The auditor was satisfied with the standard operating procedures as well as the administrative procedures.

6.1.3.2 Protocol amendments

Auditor's comments:

As mentioned earlier, the auditor thinks the trial would have benefited had DEFRA produced a project and protocol document. Over the course of the trial changes have been made, quite understandably to the way the trial has been conducted. These changes have been recorded along with justification for the change, e.g. A change to the way samples are cultured, but it would have been preferable had these changes been documented in a formal project / protocol document, along with the detailed justification for the change.

Auditor's recommendation:

*The auditor recommends that a list of definitions and a summary of changes over time, including the rationale for any changes, appears in the project and protocol document. (**recommendation 1**). Such a document will form an invaluable aid for future analysis.*

6.1.4 Communication

In general, communication between all parties involved in data collection was good. There has been good communication between the two WLUs and between the WLUs, the VLA and the ISG. In addition there have been regular data handling meetings where the ISG, DEFRA, WLU and the VLA were represented.

Nevertheless, during the visit to Polwhele WLU, the auditor came across a very detailed memo, which had been circulated to a limited number of staff, explaining how to handle the completion of the RCT 10 form. This memo was written by the VLA manager responsible for badger post mortem examinations because of the potential for ambiguity and confusion in the completion of the form. It was written to provide guidance, however it was apparent that this memo had not been circulated to all parties involved in the trial e.g., it had not been sent to the VLA data manager.

6.2 Data management

6.2.1 Data management activities

A designated trial data manager is in post at Aston Down. The WLU trial data manager has liaised with the VLA over data validation that can be added to the data entry programs, and over data queries that can be run to identify erroneous data. The VLA acts as a service in terms of data management – its main role is to implement changes or enhancements.

During the visits to Aston Down the WLU trial data manager explained his work and responsibilities. Data and cross consistency checks being undertaken include a number of checks to ensure the completeness of the data.

- check duplicate survey data
- check for incomplete badger records

In addition checks of the following datasets were carried out:

- 20% survey to update data from field staff
- 10% sett checks for illegal disturbances in control areas

The main purpose of the 20% survey check was to provide updated data on setts, latrines and signs of badger activity within triplets after defined periods of time. The

10% sett checks were to look for signs of illegal disturbance with badger setts in the control areas to provide qualitative data on possible unlawful removal of badgers in areas not subject to a badger removal treatment.

During the visit to Polwhele the Unit administrative team took the auditor through examples of data checks carried out, including their recording in the data check sheets. The auditor was impressed by the thoroughness of the process and by the quality of the documentation linked to these quality control checks.

In addition to the above, the administrative staff carried out monthly checks to clean data and the trial data manager would periodically check these. The checks that were being undertaken had been defined at the data handling meetings, and some had been requested by the ISG.

6.2.2 Data cleaning

Data cleaning was carried out during the trial as an iterative process. Nonetheless further cross consistency checks remain to be defined and carried out. As an example of the type of cross consistency checks and data cleaning that should be carried out, the auditor looked in detail at one form, the Sett Survey Form (RCT4).

Auditor's recommendation:

*All data entry forms should be examined to see if further cross-consistency checks can be identified. If such checks are identified appropriate queries should be run against the data, and appropriate corrections done allowing the data to be "cleaned" (**recommendation 2**).*

This cross consistency check carried out by the auditor highlighted also what were originally thought to be errors, in that what appeared to be 'duplicate setts' had different data associated with them. For example 125 of the 7463 'duplicate setts' had different numbers of active holes, when the auditor's assumption was that they should have the same number as they were duplicates.

The explanation given to the auditor was that despite having the same identifier and having been visited on the same day, these were not the same setts,

Auditor's comment:

The auditor had assumed, as others would have; that different setts labeled with the same identifier in different parts of the database where the same set. In fact they were not.

Auditor's comments and recommendations:

*It is very possible that future researchers looking at this data (RCT4 data) would make the same assumptions that the auditor has made. Therefore very carefully worded documentation needs to be associated with the data to avoid potential misinterpretation of the data. (**recommendation 1**).*

6.3 Key Data

6.3.1 Population data

6.3.1.1 Recruitment

The population data represent the number of occupiers within trial areas, whilst the trial population is the number of these that have been recruited into the trial. In theory both the population data and the trial population data are stored in the trial database, with the original VetNet data supplied at the start of the trial (the VetNet Occupier data) representing the population data and the occupiers that were signed up (the occupier data) being the trial population.

The ISG explained that there was no difference in terms of incidence analysis between occupiers whether or not they were recruited into the trial, and clarified the difference between a recruited and non-recruited occupier:

"All herds within trial areas (those identified as being in the eligible population) were subject to annual herd tests and their data were recorded in VetNet. These tests were performed and the data made available for analysis regardless of any agreement (or lack thereof) on the part of the cattle owners. On the other hand, badger surveying and culling could only be undertaken with the agreement of the landowner/occupier. Thus, badger culling required first identification of the relevant owner/occupier and second their consent to "surveying" or "surveying and culling"."

6.3.1.2 Identification and recruitment of occupiers

The information on potential trial occupiers was originally supplied via VetNet, and this information is held in the VetNet Occupier table. The process that should have been followed during the recruitment period was for the data on recruited farms to have been transferred from the VetNet Occupier table to the Occupier table. It should be possible to obtain the population data by reference to the VetNet Occupier table and the Occupier table.

A study of the VetNet Occupier table shows that there does not appear to have been consistent results between the WLUs' in terms of recruiting farms into the trial (see table 2). There are a far higher number of 'Not recruited' in the trial areas administered by Aston Down.

Table 2 The ratio of occupiers recruited (In the Occupier file) to not recruited (In the VetNet Occupier file)

Trial area	'Recruited'	'Not Recruited'	Ratio	WLU*
A1	251	292	1.16	A
G3	264	230	0.87	A
G2	405	352	0.87	A
A2	300	260	0.87	A
A3	181	151	0.83	A
D2	234	178	0.76	A
E3	245	173	0.71	A
I3	235	158	0.67	A
E1	248	165	0.67	A
B2	307	185	0.60	P
D1	247	137	0.55	A
D3	291	144	0.49	A
E2	227	106	0.47	A
I2	290	126	0.43	A
G1	305	132	0.43	A
F3	674	254	0.38	P
I1	195	73	0.37	A
F1	268	73	0.27	P
J2	321	85	0.26	P
H3	339	82	0.24	P
H1	201	45	0.22	P
J3	355	79	0.22	P
B3	266	59	0.22	P
J1	303	60	0.20	P
B1	206	37	0.18	P
F2	531	74	0.14	P
H2	250	34	0.14	P
C1	271	13	0.05	P
C3	264	9	0.03	P
C2	331	8	0.02	P

*: A: Aston Down, P: Polwhele

The 'not recruited' field in table 2 represents two types of occupier; First, occupiers that were found to be outside of trial areas, and secondly those within trial areas that were not recruited. Because table 2 includes details of occupiers outside trial areas, the apparent bias may, in fact, just be a result of the accuracy of the original VetNet data supplied. However, Table 3 shows the results of an analysis to ascertain occupiers within the Vetnet Occupier file that appear to be within trial areas but not recruited. This again shows a bias in recruitment between the two WLUs.

Table 3 Occupiers in the VetNet Occupier file that remained valid for selection at the time of recruitment i.e. appear never to have been attempted to have been recruited

Triplet	No. of Occupiers	WLU*
G	82	A
A	72	A
E	59	A
D	57	A
B	43	P
F	41	P
I	29	A
C	16	P
H	6	P
J	6	P

*: A: Aston Down, P: Polwhele

In following up these apparent differences in recruitment, it was found that there were different approaches used at the two WLUs. Staff at Aston Down said they had found the original VetNet lists of occupiers of little use, and had worked from maps to recruit farmers to the trial, whereas staff at Polwhele had made more use of the VetNet Occupier data. These different approaches may, to some extent, explain the differences found.

Auditor's comments:

This inconsistency between the approach to recruiting farms into the study and the apparent differing success rates, poses a potential problem although randomization provides protection against such bias.

At the meeting with the ISG statisticians they said that they were aware that in some instances the two WLUs had acted differently. For this reason they have added a possible WLU effect into their analyses but no effects were found.

6.3.1.2 Trial GIS Data

GIS data represent a further source of trial population data. Occupiers that are recruited to the trial have their land parcels digitized. Cross consistency checks have taken place between the trial database and the GIS coverages. The auditor checked 11 occupier entries from the database against the GIS, where the land distribution was not straight forward e.g. they had land in the outer buffer and in the trial area, but not in the inner buffer area. In general there was good consistency between the two data sources (table 4), with 2 occurrences of errors where the GIS data was not consistent with data held on the trial database, and one instance where the data was not there.

Auditor's recommendation:

Because the recruitment of occupiers is an ongoing process throughout the life of the trial, the process of checking the trial database against the GIS coverage should be repeated at the end of the trial (**recommendation 3**).

Furthermore the GIS data should include all the population farms, which should then be cross-referenced against the population data.

Table 4 Occupier data cross referenced between GIS and trial databases

Occupier ID	Trial Area	Treatment Area	Inner Buffer	Outer Buffer	Validation check
A1 336	P	Y	N	Y	Ok
A1 924	P	Y	N	Y	Ok
A2 971	Y	N	Y	N	Ok
D1 191	P	Y	N	Y	Ok
D2 332	Y	Y	N	Y	Ok
D3 053	Y	N	Y	N	Trial area flag set wrong on DB
E2 283	Y	N	Y	N	No data on GIS - refusal
E2 306	Y	N	Y	N	Ok
D2 332	Y	Y	N	Y	Ok
G2 631	Y	Y	N	Y	Treatment, Inner and outer flags set wrong on DB
I1 050	Y	N	Y	Y	Ok

Y = wholly within

P = partially within

N = outside

6.3.1.3 Changes in ownership or withdrawal from the trial

From the inspection of occupier paper files (source data) at the WLU it appears that in the early years of the trial, farms that i) wished to withdraw from the trial, ii) were taken over or iii) went out of business were deleted from the trial database. It is not clear how long this process went on for, but in the later years of the trial such data have been archived. The WLU informed the auditor that they would not expect there to be many occurrences of 'deleted occupiers', because change of ownership etc is quite a rare event. The VLA data manager informed the auditor that if occupiers had been lost, data on other tables, surveys particularly, would be isolated, which is not the case. Where that did occur, the deleted occupiers were put back on request, but these were few. Other reasons for deletion were that they were virtually duplicate records and two records for the same occupier were unnecessary.

For the analysis of the data, using occupiers as defined by the trial database, deletion of such data represents a problem, primarily because if any of these 'deleted' occupiers had a TB breakdown during the period before deletion, this information will not be available for analysis. This is because data on TB breakdowns

are obtained by linking the occupier data to VetNet TB breakdown data via the identifier of the farm, which will no longer exist.

Auditor's comments:

Data on changes to farm details, such as dates when farms were archived, exist on VetNet. An attempt should be made to re-create the details of these 'deleted' farms (recommendation 4).

6.3.2 Herd Breakdown data

6.3.2.1 Collection of breakdown data

The number of breakdowns in trial areas is currently obtained by linking the Trial occupier data to details of breakdowns held on DEFRA's TB disease database (VetNet). This presents a problem, however, because occupier data within the trial database is recorded at the level of County/Parish/Holding (CPH), whilst breakdowns on VetNet are recorded at the level of County/Parish/Holding/Herd (CPHH). Hence, if an occupier exists with land both within and outside the trial (which may be the case for both single or multi herd occupiers), this method can not distinguish where the breakdown occurred.

Occupiers owning land both within and outside the trial are likely to be a particular problem for multi-herd occupiers. Although the number of such multi-herd occupiers is quite low (for the 3872 occupiers where CPH is recorded, 91 (2.4%) have more than one live herd, with a maximum of 4 live herds). The way in which breakdown data are recorded on VetNet does not adequately handle such situations as rented land and the movement of herds over time and some doubts remain over the location of such breakdown data, however attempts have been made to ascertain where breakdowns occurred for such occupiers

Ambiguities arising from the recording of CPH rather than CPHH have been resolved by linking the RBCT Occupier number and the VetNet data for the herd (tests, breakdowns, etc). This link is made by using a table within the trial database constructed and maintained by VLA that lists for each occupier number the VetNet core id (or core reference number) for the herd. The core id, like CPHH, defines distinct herds within a holding and should prevent the incorrect attribution of breakdowns that occurred on an unrelated herd to a herd enrolled in the RBCT that has the same CPH.

Auditor's comments

The best method of obtaining trial breakdown data would have been to capture this data directly at the time of the breakdown.

Any incorrect attribution of breakdowns should be overcome by using the link table described above.

6.3.2.2 RCT 17 Data

The RCT 17 form should record breakdowns within trial areas, even though it does not form part of the trial database. The SVS is responsible for generating RCT 17 forms. It appears that RCT 17 forms have not been generated for all breakdowns in trial areas but rather for the purpose of:

1. Informing the VLA of the need to issue controls for the TB99 study. The TB99 study involved all trial areas only up until the end of 2003; therefore a definitive list of breakdowns can not be obtained from this study.
2. For informing the WLU of a breakdown in a reactive trial area (which was suspended in November 2003).

6.3.2.3 AHDO data on breakdowns in trial areas

Details of what breakdowns occurred within trial areas have been kept on local spreadsheets by the AHDOs.

A validation exercise was performed by comparing the breakdowns recorded within trial areas by two AHDOs (Taunton and Truro) in 2004 against the data that would be obtained by linking the Trial and VetNet data, (see table 5).

Table 5 Validation of Trial breakdowns against local AHDO data

All breakdowns in trial areas 2004	Original local data	Original Occupier/Vetnet data	No. in local data not in occupier/VetNet (after review)	No. in Occupier/VetNet not in local data (after review)
Taunton AHDO	30	21	11 (1)	2 (1)
Truro AHDO	124	135	9 (2)	20 (12)

The original data supplied by the AHDOs showed quite large differences in the number of trial breakdowns. The auditor then asked the WLUs, VLA and the two AHDOs to review the differences. Interestingly the majority of the original differences were where the local data had relied on the VetNet identification of trial occupiers on VetNet using the 'Vetnet trial indicator'.

What table 5 shows is that, after reviewing the data, 1 breakdown should be added to the breakdown list for Taunton and two for Truro for 2004, and that local staff believe only one of the two extra breakdowns for Taunton is correct and only 12 of the extra 20 Truro breakdowns are correct.

The conclusion of the comparison of Trial/VetNet breakdown against the source breakdown data is that some differences do occur. These differences need to be investigated and a definitive list of breakdowns in trial areas compiled.

A second important conclusion is the danger of relying on VetNet location data.

The Vetnet location data originally used by the AHDOs to identify whether a farm is in a trial area (a 'trial identifier' flag within VetNet) has not been used in any part of the analysis, but the errors in the VetNet Occupier data must pose the question as to the reliability of VetNet data for other parts of the analysis.

Auditor's recommendation:

*There needs to be a validation exercise undertaken that ensures that all breakdowns obtained by linking to VetNet are truly within trial areas. When this list of "validated" breakdowns is obtained it should form part of the trial database (**recommendation 5**).*

Auditor's comments:

However the fact that this was not done does not compromise the trial's primary conclusions.

6.3.3 Badger TB prevalence data

The *M. bovis* status of badgers is defined by their culture result, PM examination and histopathology result. Source data for these results does exist within the VLA which performs these tests and examinations. Furthermore, the VLA records the results onto its own Laboratory system.

A number of queries were run, identifying discrepancies between the Trial database results and those on the VLA system. The discrepancies were investigated by going back to the source data. In nearly all cases the data as recorded on the Trial database were correct, with the VLA data being erroneous.

In addition to the standard culture work, a certain proportion of badgers underwent an extended PM and culture procedure as part of a research project. These data do not form part of the Trial database. However, concerns were raised, at the WLU, that some 'extended culture' results had been sent to them in error. Identification of whether a result was 'standard' or 'extended' was not easy because the same culture result form was used in both cases.

Auditor's recommendation:

*The auditor has not had time to investigate this further, but because badger prevalence data will be crucial for future analysis, an audit of the extended culture work forms one of the recommendations of this report (**recommendation 6**).*

6.4 ISG analysis

The ISG has undertaken two parallel sets of analyses. One based on farm population data obtained from VetNet recorded locations, and one based on the trial database.

6.4.1 Extraction of trial database data

To validate the data that the ISG extracted from the trial database for their proactive analysis, the auditor asked the VLA to reproduce a sample of the dataset blindly, using definitions supplied by the ISG, but using their own methods. Table 6 gives the results of this comparison.

Table 6 Comparison of data extract used for analysis between the ISG and VLA

Triplet	Treatment	ISG Confirmed Breakdowns	ISG Baseline herds	VLA Confirmed Breakdowns	VLA Baseline herds	ISG - VLA Breakdown difference	ISG - VLA Baseline Herds difference (%)
A	Proactive	38	59	38	64	0	-8.5
A	Survey-only	55	75	56	79	-1	-5.3
B	Proactive	91	132	91	135	0	-2.3
B	Survey-only	61	112	61	117	0	-4.5
C	Proactive	36	104	36	107	0	-2.9
C	Survey-only	93	169	93	175	0	-3.6
D	Proactive	29	67	29	75	0	-11.9
D	Survey-only	36	69	37	76	-1	-10.1
E	Proactive	40	86	40	88	0	-2.3
E	Survey-only	50	79	50	79	0	0.0
F	Proactive	15	111	15	117	0	-5.4
F	Survey-only	70	183	70	188	0	-2.7
G	Proactive	65	173	65	182	0	-5.2
G	Survey-only	40	103	40	113	0	-9.7
H	Proactive	34	55	34	58	0	-5.5
H	Survey-only	32	117	32	122	0	-4.3
I	Proactive	30	82	30	85	0	-3.7
I	Survey-only	19	78	19	81	0	-3.8
J	Proactive	39	117	39	n/a	0	n/a
J	Survey-only	37	122	37	n/a	0	n/a

Table 6 shows a very close agreement between the VLA and ISG for the number of confirmed breakdowns, with there only being a difference of two over the whole life of the trial.

The differences for the baseline herd figures were discussed at the meeting with members of the ISG statistical group. It transpired that the ISG had made an assumption that if a herd had not been tested over the life of the trial, it did not in fact exist, even though it was marked as 'live' on VetNet. The VLA investigated a sample of these occupiers, by referring back to the AHDOs, and for the majority of occupiers the assumption made by the ISG is correct.

The error is most likely to lie with the VetNet data, in that these herds have not been marked as no longer trading (or archived)

The VLA data manager informed the auditor that these herds are likely to remain as 'live' on VetNet for many years with no stock if the owner doesn't declare stock will not be reintroduced at some point.

7. Auditor's recommendations

7.1 **Recommendation 1:** Project and protocol document (see 6.1.2.1, 6.1.3.2, 6.2.2)

In order to aid future analysis by individuals who have not been involved with the trial data, I recommend that a project and protocol document that contains details and definitions of key data and an explanation of how they can be recreated from the raw data is drawn up: e.g. the definition of a trial breakdown. This document should also include a summary of the changes over time. Such a document will form an invaluable aid for future analysis.

7.2 **Recommendation 2:** Database "cleaning" (see 6.2.2)

All data entry forms should be examined to see if further cross-consistency checks can be identified. If such checks are identified appropriate queries should be run against the data, allowing the data to be cleaned. It is essential that the data is as clean as possible for future usage, particularly at the point it is opened up to 'outside' analysis.

This work has now been completed and checked by the auditor.

7.3 **Recommendation 3:** GIS database (see 6.3.1.2)

There should be a final process of checking the GIS database against the trial database for cross-consistency in terms of trial occupiers. Furthermore the GIS database should be extended to include all occupiers within trial areas, whether or not they had been signed up to the trial. This process would provide validation of the trial population figures.

7.4 **Recommendation 4:** Changes in ownership or withdrawal (see 6.3.1.3)

As data on changes to farm details, such as dates when farms were archived, exists on VetNet, an attempt should be made to recreate occupiers that were deleted from the trial database before the process of archiving occupiers was implemented.

This work has now been completed.

7.5 **Recommendation 5:** AHDO data on breakdowns in trial area (see 6.3.2.3)

There should be validation of the 'within trial area breakdown list' against data held locally at the local AHDOs. When a validated list of breakdown list is compiled it should form part of the trial database, rather than having to be recreated for each new analysis. The breakdown list should be updated at regular intervals.

7.6 Recommendation 6: Badger TB prevalence data (see 6.3.3)

For all badgers that formed part of the 'extended culture' work there should be a check that the culture result that exists on the trial database, is the culture result according to the standard protocol. This is because of concerns raised at the WLUs that some of the culture results forwarded to them for data entry may in fact have been the extended culture results.

This work has now been completed and checked by the auditor.

Further recommendations not previously made in the report

In future, audit activities should best be initiated at the planning stages of a trial rather than when a trial is already well in progress.

At some stage DEFRA will have a responsibility to "sign off" the data, after which no changes to the data can be made.

DEFRA has also a responsibility to ensure that the data, including source data, are archived in such a way that they can be retrieved at short notice.

DEFRA needs to maintain control over the future release of the data. It should consider forming a review panel, that includes those who have been responsible for the data, to look at any communications prior to publication, that result from the release of this data.

References:

- ¹ J.R. Krebs, R. Anderson, T. Clutton-Brock, W.I. Morrison, D.B. Young, C.A. Donnelly, S.Frost and R. Woodroffe. Bovine tuberculosis in cattle and badgers – An independent scientific review. PB3423, 1997 MAFF (London).
- ² Bourne J, Donnelly C, Cox D, Gettinby G, Morrison I, Woodroffe R and McInerney J. Towards a sustainable policy to control TB in cattle: a scientific initiative - First report to the Rt Hon Dr Jack Cunningham MP from the Independent Scientific Group on Cattle TB, Ministry of Agriculture, Fisheries and Food, 45pp, 1998.
- ³ Mollison, D. First Report of the Statistical Auditor on the badger culling trial. Defra Publications 2000, PB 5385. (This report, and two further reports published in 2004 and 2005, is available at: <http://www.defra.gov.uk/animalh/tb/culling/p5aud.htm> .)
- ⁴ H.C.J. Godfray, R.N. Curnow, C. Dye, D. Pfeiffer, W. J. Sutherland and M.E.J. Woolhouse. Independent scientific review of the randomized badger culling trial and associated epidemiological research, 2004.
(Available at: <http://www.defra.gov.uk/animalh/tb/default.htm> .)
- ⁵ Woodroffe R, Donnelly CA, Johnston WT, Bourne FJ, Cheeseman CL, Clifton-Hadley RS, Cox DR, Gettinby G, Hewinson RG, Le Fevre AM, McInerney JP and Morrison WI. Spatial association of *Mycobacterium bovis* infection in cattle and badgers (*Meles meles*). *Journal of Applied Ecology* 42, 852-862, 2005.
- ⁶ Woodroffe R, Donnelly CA, Jenkins HE, Johnston WT, Cox DR, Bourne FJ, Cheeseman CL, Delahay RJ, Clifton-Hadley RS, Gettinby G, Gilks P, Hewinson RG, McInerney JP and Morrison WI. Culling and cattle controls influence tuberculosis risk for badgers. *Proceedings of the National Academy of Science* 103, 14713-14717, 2006.