Chapter Seven: Identification of the Soldiers Buried at Pheasant Wood, Fromelles

by Margaret Cox, Peter Jones, Louise Loe and Richard Wright

SUMMARY

The principal aim of the Fromelles Project was to identify as many as possible of the soldiers whose physical remains were excavated from the graves adjacent to Pheasant Wood. The rationale for the anthropological and artefactual analyses has been to provide a suite of post-mortem data to compare to the ante-mortem data of the missing soldiers as deduced from historical sources. While the archaeology and anthropology have been detailed in chapters one to six, no mention has yet been made of the work of a group of researchers whose principal role has been to track down as many of the families of the missing soldiers as possible and ask them, or another family member, to donate DNA, if they are an appropriate donor. The ante- and post-mortem datasets gathered by archaeologists, anthropologists, molecular geneticists, historians and genealogists, plus the DNA profiles of the missing soldiers as indicated by their families, provide the materials needed to attempt to attribute identification to the buried soldiers, where this is possible.

As no similar sized project of this type has been undertaken before, it was necessary to establish a methodology for establishing identifications; further, a decision had to be made as to the level of proof required before an identification could be attributed. For missing persons of almost one hundred years ago (and older), the use of DNA extracted from such people can only utilise inherited markers associated with the maternal and paternal line, which inevitably means lower levels of match probability when compared with autosomal DNA analysis associated with scene of crime investigations. As such, it is not reliable or ethical to attribute a name at a level that is 'beyond reasonable doubt'. However, our datasets do facilitate the attribution of identification (ID) at a level where 'clear and convincing evidence indicates that an ID is substantially more likely than not'. The distinction between the two levels of proof is important, as it conveys the reality of the evidence and not an aspirational level that is inappropriate.

The ID process developed by us for use on this project is one that is rigorous, repeatable and free from bias, one that makes use of relevant statistical tests, and one that ideally moves from the most heavily weighted datasets to those that are less discriminatory. While DNA is the prime mover in this process, it comes with limitations, and an ID indicated by DNA is never uncritically accepted where other datasets strongly suggest that the indicated ID is not correct. Our maxim for attributing an ID is that there must be 'at least three legs on the stool', that an ID is strongly suggested by at least three datasets with no further dataset being contradictory.

At the time of writing, we have attributed names to 144 of the missing soldiers from the Battle of Fromelles, and are presently working hard to trace additional donor families and other data for soldiers who currently remain 'known unto God'. Although the Fromelles Management Board will cease to exist after 2014, each country will then assume responsibility for subsequent analyses and identifications.

INTRODUCTION

Human identification is a complex issue and 'identification' is a term with several meanings. There are two types of ID that are relevant to this project. The first is social or cultural ID. This is the name by which an individual is known in life and all that is implied by that name, including assumed genetic relationships to others in the same family group, and an individual's place within society. If we were aiming to attempt to ID the soldiers buried in the mass graves at Pheasant Wood using data other than molecular in character, then the type of ID we might achieve would be social, or the name by which a soldier was known during his life. Of course, for the majority, this ID is also their genetic ID. However, with the use of DNA profiling a shift takes place, especially where the genetic information is heavily weighted in the ID process, and where it is often the most statistically robust data available. Here, we are looking at 'genetic ID', and this may or may not be the same as 'social ID'.

It is a given in all cultures and at all times that non-paternity and adoption take place, both within and without wider family groups, and that the rates of both are neither constant nor particularly predictable. A typical estimate for non-paternity is quoted as 5-10% (Gilding 2005), but this figure has been revised down to 1.9% dependent on the population and cultural group (Anderson 2006). There are no figures for the rates of non-paternity from the early 20th century, and our working assumption is that it will be around 1-2%. There is an inherent loss of a number of IDs, because some individuals, who may have been either adopted or illegitimate, will not be identified to their 'social family' group because they do not share a genetic background with their presumed descendants. Regrettably, because their existence is usually unknown by genetic relatives alive today, they will remain forever unidentified. Such is the cost of a greater degree of certainty in identifying the majority of the dead utilising molecular genetics.

This issue is one that raises ethical considerations that do not arise when other positive ID characteristics of the missing are recorded during life, and are available for use today, such as fingerprints, or accurate dental records that can be compared with post-mortem data derived from surviving fingerprints or dentition extant at death. How to deal with incidental findings (IFs) arising during an ID process is a complex issue and must be agreed at the outset. The literature on this complex issue begins around the time that this project commenced (2008); for details see Wolf *et al.* 2008; Parker 2008, and Parker *et al.* 2013. IFs may also highlight genetically related individuals amongst the Missing to donors, of whom the surviving family has no knowledge because their existence was concealed at the time of their birth and during their and their parents' lifetimes. Such individuals, lacking DNA associated with their social family, will not be identified to the name by which they were known in life, that is, their social family groups, unless there are exceptional and over-riding biological datasets that are powerful enough to overrule the genetic indicators. Where multiple same-sex DNA donors from a family are employed, as has sometimes been the case at Fromelles, this too has risks as it may reveal that donors in the same social family group may not be related in the manner they have always presumed.

The issue of whether or not to disclose IFs to relevant families was an ethical consideration for the Fromelles Project. However, it was rendered less of a dilemma because of the nature of the project and the confidentiality of the proceedings (reflecting its official military and consequently confidential nature), whereby the buried soldiers were treated in the same manner as our war dead who are killed in action (KIA) today. In line with the recommendations of Parker *et al.* (2013), donors are not informed where this occurs because to do so could negate the potential benefits of the process. Recognizing that the risks of discovering kinship discrepancies are not uniform across space and time (ibid., 225), we consider that each case should be considered on its individual merits. Exceptions to a policy of non-disclosure might include projects where the potential for substantial psychological or material benefits outweigh the risks as in more recent mass fatality incidents (ibid., 224).

All the work undertaken for this project was fully compliant with all relevant Australian, UK and EU regulations and legislation, and adhered to all international guidelines relevant to the different aspects involved. A number of guidelines have been produced by the UN, governments and NGOs and these include those produced by the Inforce Foundation – Cox *et al.* (2008), the US Department of Justice Office of Justice Programs (National Institute of Justice and US Department of Justice 2005), Interpol (2009), and the International Committee of the Red Cross (2009). All stress the message that forensic scientists (and those involved in identification programmes in other capacities) must recognize that they have an obligation not only to the legal (and other) institutions that retain their services, but of equal importance, to the families of the Missing (Stover and Shigenake 2002, 864).

Identification of the dead is discussed in depth in Thompson and Black (2007) and the reader is referred to that text for further information. Post-mortem ID relies upon a combination of available datasets, comprising ante-mortem data and post-mortem data. Ante-mortem data are what can be established about an individual's life up to the point of their death (usually documentary/historical/medical records etc.). Post-mortem data are information retrieved about a person's life and death from their body, including artefacts buried with them. At the outset of the ID process we had ante-mortem data (mainly historical sources) for most of the missing soldiers (n=*c* 1650) and post-mortem data for the buried soldiers (n=250).

Two levels of ID are relevant to this project:

i. Positive ID. This is achievable when unique ante- and post-mortem data types (fingerprints, dental records, DNA) match. Other biological evidence must support this ID. (For Fromelles we only have DNA.) If these data contradict the ante-mortem data for the individual under consideration, then issues of non-paternity and other irregularities are considered but in historic cases these can rarely be resolved.

ii. Presumptive ID. This is achievable when no single factor alone justifies a positive ID, but taken together all the factors strongly suggest that the postulated ID is very likely to be correct. Such factors might include a number of skeletal indicators such as unusual stature estimate (short or tall for the cohort) and age at death estimate (young or old for the cohort), handedness, evidence of a known (from medical records) healed fracture, racial characteristics, a DNA match with a low probability level, plus some artefactual evidence. DNA haplotypes may be useful in cases where they indicate a particular and unusual geographical region or ancestry, especially in cases where the anthropology also suggests a particular ethnic origin.

THE IDENTIFICATION PROCESS

The methods used for identifying the buried soldiers have been discussed previously by Cox (2010) and Jones (2010), and what follows is modified and enhanced from that. It is important at the outset to appreciate that our efforts to identify the buried soldiers from Pheasant Wood are not the same as a 'normal' archaeological or historical project where there should be complete transparency in methodology and detail. This reflects that this project is concerned with both British and Australian war dead, many of whom have living relatives (including some siblings) and who are accorded the same dignity, respect and privacy as any currently serving member of the armed forces. Consequently, the process developed for this project is compatible with the Australian Defence Forces (ADF) policy for ID of their war dead and is compatible with the principles, ethics and philosophy outlined in the Interpol DVI Guide (Interpol 2009). Information management and confidentiality issues are therefore the same as for modern war casualties.

As outlined in Chapter One, a Data Analysis Team (DAT) was formed at the outset of this project (2009) to conduct the ID process. The DAT comprised a chair and deputy chair, and subject matter experts in anthropology, archaeology, molecular genetics and military history and records. The chair and deputy chair played dual roles as they are also subject matter experts. They were not, though, the sole experts assisting the DAT in their fields. Other experts, such as a statistician, were consulted as necessary. The function of this team was to annually analyse all available datasets in a systematic manner leading towards placing each of the buried soldiers in one of the ID categories described in Table 7.1. The DAT's deliberations concluded with recommendations being made to the Joint Identification Board (JIB) for their consideration. Project administrators kept records and the whole process was recorded for ADF protocol, as it was effectively a Board of Inquiry. The process also benefited from the advice and input of an experienced external quality assurance expert (Alison Taylor, HM Coroner, who has extensive experience of dealing with mass fatality incidents in the UK and elsewhere). Her guidance and impartiality were both helpful and welcome, and a measure of good practice in this process. As no project similar to Fromelles has previously been undertaken, it was necessary to develop a protocol for the ID process at the outset; one based on available datasets, their inherent potential and limitations.

ID categories

There were three possible levels of ID that could potentially be attributed to the buried soldiers from Fromelles and it was considered imperative that the process and the weight of evidence for each was stringent but achievable with the available datasets. After due consideration and external advice it was considered impossible to identify a buried soldier to a name at the level of proof 'beyond reasonable doubt' with the datasets available to us. The level of 'clear and convincing evidence that indicates that an ID was substantially more likely than not' was, however, considered to be achievable, and this standard has been used for IDs in Australia previously. It was agreed at the outset that this would be the level of confidence used when attributing a name to any of the buried soldiers excavated from the graves at Fromelles. To identify a soldier to the army for which he fought was considered something that could reasonably be attributed at the level of 'balance of probability'. That a soldier fought for any particular army could not however be seen as evidence that they were of that nationality. Where no such attribution was possible, a soldier's ID remained 'unknown'.

For this process to reach its ultimate aim, that is, attributing a level of ID to each of the buried soldiers, there had to be a hierarchy of decision-making. The decision to attribute a name to one of the buried soldiers (Category 1 ID) remained the sole responsibility of the national member(s) representing the government of the force the soldier served at the JIB, which sat every year from 2010 to 2014 (since 2014 subsequent identifications have reverted to being a national responsibility). In order to reach a decision the JIB considered and scrutinised the recommendations of the DAT. The relevant national member consulted with his/her JIB colleague during the

Table 7.1: ID categories

Category	*Description*	*Level of proof*
Category 1	Soldiers identified by name	Where there is 'clear and convincing evidence that indicates that an ID is substantially more likely than not'
Category 2	Soldiers identified only by army, or army or regiment/battalion and/or rank	Where sufficient evidence exists to assign an army at the level of 'balance of probability'
Category 3	Soldiers deemed to be unknown	Where insufficient evidence exists to place a soldier in a higher category

decision-making process but retained the final responsibility for making such a decision.

The decision to place a buried soldier in Category 2 or Category 3 had to be agreed by both the UK and Australian Members of the JIB. Artefactual evidence was the main source of evidence that identifies one of the buried soldiers to Category 2 with, in exceptional cases, exclusive ethnic ID (for example, Australian Aboriginal) also being considered. Placement in either category was a judgement based on the balance of probability, and having both members of the JIB agree strengthened the veracity of the decision. If a consensus for a Category 2 decision was not reached, then the soldier was ascribed as Category 3.

All evidence pertaining to the ID of the 250 buried soldiers was considered on a case-by-case basis. A consistent holistic approach to reviewing evidence was undertaken to ensure that all forms of evidence were examined and cross-referenced enabling a full profile of each case to be generated before making a recommendation to the JIB. Decisions and recommendations were recorded, together with details of all of the evidence supporting them, and any qualifications, dissenting views, or contra-indicative evidence. A bespoke *pro forma* body summary data sheet was used to ensure that all such detail was recorded and reviewed each time a case was considered.

Clearly, standards of evidence were critical in this process and were both a professional and an ethical consideration. As noted above, Category 1 decisions were based on a suite of clear and convincing evidence that, where possible, included a DNA match and indicated that an ID was substantially more likely than not. No evidence type was used exclusively. For each soldier, we examined and considered all available datasets. In reaching a decision to name a soldier, there had to be an absence of relevant unexplainable or compelling contradictory evidence. In practice, this meant that reaching some Category 1 decisions could be delayed while additional analysis and research was undertaken and discrepancies resolved where possible.

Assignment of a Category 2 ID was based on the balance of probability and this in turn rested upon the weight of evidence. The DAT formulated an assessment framework (Table 7.2) that describes how different evidence was weighted. The level of weighting in Table 2 is qualitative, as we could not find a suitable and reliable method of attributing values quantitatively, despite initial efforts.

The process used by the DAT to assign an ID to the buried soldiers was identical for each case. Ideally, subject to all datasets being available at the outset (including DNA from the missing soldiers' families), this moved from the results of DNA analysis to consideration of all post-mortem data both in their own terms and in respect of ante-mortem data for potential candidates derived from the list of missing soldiers. The statistical program OPTIMISE (see below) was used as an aid at this point and was based on nearest neighbour analysis of age and stature. Both the distance and rank indicated by the method were recorded, the strongest value for distance being naught, signifying identical age and stature. To facilitate this top-down process, once the buried soldiers' DNA profiles were obtained, then each buried soldier was assigned one of three DNA categories:

DNA Group I - Matched profiles:

The buried soldiers in DNA Group I were those whose profiles had the following characteristics:

1. They matched a donor family's DNA for either the Y or mitochondrial profile or both, but in all cases with a strong match probability.
ii. The profile was not present in the Elimination Database.[1]
iii. The profiles from the missing soldiers via their informative family members were consistent and the potential for non-paternity and adoption was eliminated.

DNA Group II - Unmatched profiles:

Those buried soldiers whose profile did not match a family currently held on the DNA donor database, or was matched to a number of families. This may reflect partial profiles, common mitochondrial or Y profiles, or contradictory profiles; that is, the relevant Y and mitochondrial profiles did not match suggesting non-paternity or non-maternity within a family. Cases from this group fell into one or more of the following characteristics:

i. DNA was profiled for the buried soldier but did not match any family currently held on the DNA donor database.
ii. Multi-matching: As per Group I but the profile of the individual matched more than one family group.
iii. The profile matched a particular haplogroup.
iv. Cases of suspected non-paternity and adoption.

The third DNA group was to have been those buried soldiers that did not yield a DNA profile. However, this group allocation was not utilised.

Those cases with a unique DNA match (DNA Group I) were considered first, followed by those from Group II. Profiles from donors that were the

[1] To ensure that no cross contamination of DNA samples took place during the excavation and analysis process, all of those involved in the field and DNA laboratory gave a DNA sample to contribute towards the creation of the 'Elimination Database'. All issues of security and informed consent were complied with.

Table 7.2: Values for each type of biological and artefactual data reflecting the accuracy of the methodologies available and the confidence level associated with its context

Type of analysis	*Weighting for ID (high, medium, low or nil)*	*Weighting for army (high, medium, low or nil)*
DNA	Low to high depending upon match probabilities	High (where derived from ID) or moderate if derived from haplotypes only
Age estimation	High up to 21 years and the years around 30, but moderate otherwise, low for the over 50s range can be useful	Nil
Estimation of ancestry	Low for Caucasoid (except as exclusion) but could be medium in the case of a rare group	Low or medium but the latter only if Australian Aboriginals found
Estimation of ancestry based on DNA	Could be high in the case of rare or unusual genotypes, otherwise low	High for rare and confident army specific markers
Stature	Medium, range is useful	Nil
Handedness	Low, as can be culturally defined	Nil
Facial attributes	Nil, unless there is a good photograph that shows adult appearance; if there is then low to medium depending on degree of unusualness of the face in question	Nil unless unusual ancestry specific to only one army pertains, and in such cases low to moderate
Skeletal constitution	Nil, unless occupation and lifestyle are known. In such cases can be low to moderate if the subject had had either a very physical or sedentary occupation/ lifestyle for a considerable time	Nil
Photographic superimposition	Nil, unless the missing person has good facial photo-graph. If he does then low to moderate depending upon the degree of unusualness of the face in question	Nil
Peri-mortem trauma	Medium where records exist	Nil
Ante-mortem injuries	Low or medium depending upon the nature of the condition and its rarity	Nil
Disease	Low or medium depending on rarity	Nil
Dental health and treatment including prosthetics	Medium for Australians as dental health was recorded at enlistment and when treatment occurred while in theatre	Nil
Life style indicators – cortical remodelling	Low, too many other variables influence this	Nil
Type of Artefact - assumes confident contextual association	*Weighting for ID (not allowing for contextual weighting as outlined in chapter)*	*Weighting for army*
ID tags	High if context is secure and position is credible	High as derived from ID
Other military accoutrements (eg badges)	Low where portable	High (if not portable), e.g. jacket belt buckles
Personal items with name or initials (eg cutlery sets, toothbrushes)	Medium depending upon level of association	High if derived from ID
Footwear and fabric items of uniform	Low	Depends upon position, context and portability
Paper or card-based materials (cigarettes packets, letters, photographs etc.)	Any depending on type, rarity and country of origin specificity of artefact	Any as for ID
Material containing DNA correlated to remains (a last resort only to be used if no DNA informative relative exists from the family)	High if association is secure	High as for ID
Ante-mortem data derived from written sources	*Weighting for ID*	*Weighting for army*
Enlistment, burial and other military records	High if the soldier and or other reporting person told the truth – treat with caution and always see when the data was written in relation to when it occurred	High but treat as for ID
Questionnaire data	Will range from low to high depending on how the relative derived the information – treat with caution	High where it relates to an ID

same were grouped together. These were usually mitochondrial profiles and matched a number of bodies with similar common profiles. For each buried soldier, all datasets and evidence types were considered and assessed one against the other. The decision as to which data type was most highly weighted reflected the numerical values set out in Table 7.1. However, consideration of such issues as contextual security and the probability values of the data type may have caused deviation from such values and where this occurred it was recorded. While those individuals from DNA Group I already had provisional names linked to them, and by default such information as to which army the soldier served, all other datasets were scrutinised and assessed against what was inferred about that potential candidate (i.e. the missing soldier) from written records. Unless the datasets agreed with the result suggested by DNA analysis, then that case was subject to further scrutiny and possibly further analysis.

Initial assessment

For reasons outside our control, the DNA results from donor families (missing soldiers) were not available when the DAT first sat in 2010. As such, an initial assessment of all available data was undertaken for each of the buried soldiers. All archaeological, anthropological and ante-mortem and post-mortem data was considered. From that initial assessment no clear or convincing evidence on its own could be used to confidently assign an ID. However, in three cases items found on the body were highly significant in terms of inscriptions and their location on the body, and these might have been sufficient for a tentative identification without DNA evidence. These items were either identification discs, non-portable items with personal inscriptions, or archival information, but taken alone with our standard of proof they were not reliable enough to positively assign identification. At this stage, and lacking missing soldiers' DNA data, we made extensive use of nearest neighbour analysis to assess recorded stature and age at death for the buried soldiers against that recorded in the records for the *c* 1650 missing soldiers. This was particularly useful in cases where either age or stature were unusual, and especially so if both were. A note was made of these possible identifications in the body summary data sheet. This initial assessment was not the process that had been planned, and in subsequent identification sessions each case review was instigated by a DNA match, as per the original plan (below). This was either a Group I match or a Group II match, each of which was then further examined using all available data towards classifying the recovered remains as either ID categories I, II or III.

DNA Group I

When all necessary datasets were available then the first process for each case was to correlate and summarise that data (DNA, anthropology, archaeology, historical). This examination either supported or refuted the hypothesis that the ID suggested by the DNA analysis was correct. No attempt, however, was made at this stage to analyse the combined data. Following this, the summary information for each subset of data was checked for consistency and logical comparison of the available datasets took place before an initial assignment of the remains to ID Category 1, 2 or 3 was made. Finally, we consolidated the results into a provisional ID Category. The assignment of an ID category indicated that the DAT considered that this could be recommended to the JIB at the level of proof required for a ID category (see Table 7.1), that all information available pointed to a particular ID category, and that there were no data that substantially contradicted this view. Where an unconfirmed assignment arose, this indicated that the DAT was not confident that all information available pointed to a particular ID category at this level of proof. Where possible, further work (for example, tracking down a contra DNA donor) was conducted by the DAT (and others such as genealogists, historians, etc.), unless it had been agreed by them that the contradictions in evidence were unlikely to be resolved. While some cases from this DNA Group were recommended to the JIB for ID Category 1, others may have been placed into categories 2 or 3.

DNA Group II

Those individuals from DNA Group II may have had a number of possible provisional names linked to them, or simply evidence suggesting the army for which they fought. Assessment of these relied very heavily on post-mortem data from each of the buried soldiers, examined *vis-à-vis* ante-mortem data derived for the missing soldiers, to indicate which if any of the possible names was the correct one. The process followed was the same as for DNA Group I.

For those in DNA Group II for whom it was not possible to recommend a Category 1 ID, the next level of ID to consider was if they could be assigned to a specific army (and if possible to a battalion/regiment or division, or rank). This utilised all available archaeological evidence. In very rare cases, the buried soldier's DNA may have indicated his ancestry, and this might have suggested someone who was only likely to have served for one army or the other. While some cases from this DNA group were recommended to ID Category 1, most were placed into categories 2 or 3.

Sampling bias

It was important to deal with the potential for 'sampling bias' in this process. Bias can reflect a real trend (for example, all the soldiers from one force being buried together in one grave) or a belief or preconception (for example, the first to be recovered from the battlefield were buried in the first grave and as such they were probably all Australian, due to its proximity to the Australian positions). It was very easy to be influenced by bias in any decision-making process and therefore it was essential that any potential bias was designed out. This was helped by the allocation of body numbers as undertaken by the archaeological process as this did not indicate which grave a soldier was buried in or anything else about them. To further eliminate bias, the cases were considered in a random order. Each buried soldier's 'body number' was assigned a random number using the RAND() function in Excel. This was then sorted in assenting order and the corresponding cases when presented were considered in the order in which they appeared in the Random Body Order List. The DAT were unaware of which grave the body was recovered from, or its position in the grave, unless specific matters relating to that information were requested.

To counter any bias a random selection of 25 cases from both cohort I and cohort II was reviewed at the 2014 DAT to ensure that no bias had crept in during the DAT process. If bias was detected, those results affected were to be specifically reviewed. This process also served another useful purpose, namely to reveal if there had been any shift in the application of judgement and standard of proof through the data analysis process. If there had been, cases were to be reviewed as appropriate. At the end of the 2014 DAT we were allocated sufficient time to review all Category 1 IDs for consistency of standards and to undertake the random review of Category 2s, as described above.

Security and confidentiality

The deliberations and recommendations of the DAT were confidential and protected by rigorous security protocols and procedures. DAT members were required to respect a high level of information security and agreed not to disclose any details unless required to do so for official purposes.

Notifications

Once the DAT met and made recommendations to the JIB each year, and the JIB made its decisions, the names of those identified were checked by the Commonwealth War Graves Commission (CWGC) against its detailed database to ensure that the person identified did not already have a grave. This happened in one instance and upon examination of the basis of the original name allocation, it was considered that this had been incorrect as it had been based on very scant anecdotal evidence that would not be considered appropriate today. The list of verified named soldiers was then passed onto the relevant government ministers, who immediately thereafter advised relevant family members that their lost soldier has been identified. The families were then offered the opportunity to select the words that they would like to be inscribed onto the headstone. After families had been informed, then the media was notified. The CWGC then started the process of making new headstones for the graves affected in preparation for a rededication ceremony held each year in Fromelles on the anniversary of the battle. The residents of Fromelles are highly respectful of the sacrifice that took place in their village and its place in history, and consequently the event is always well attended with representatives from the village, local and regional officials and usually dignitaries from both governments.

See Figure 7.1 for an overview of the ID procedure.

THE DATASETS

For this project, we utilised all available and relevant methodologies that could contribute towards establishing the genetic ID of the buried soldiers. There were a very small number of cases with no possible genetic ID (for example, the family had become extinct, or no informative relatives survived, or were willing to donate DNA), where exceptional data of other types had suggested that a category 1 ID could be recommended to the required level of proof. It is worth noting that we did not use isotope analysis for this project (extracted from bone, teeth or hair) to determine provenance, as this method relies upon several assumptions that do not work on this occasion. Firstly, there has to be a good understanding of the range and distribution of isotopes across the geographical areas in question. While this currently exists for most of the UK, it does not for Australia. Secondly, the Australians present among the 250 soldiers would need to have spent a considerable part of their childhoods in Australia, which was not always the case. Research has shown that many of those enlisting in the Australian army were born in the UK (or elsewhere), and that others spent some or all of their childhoods in the UK (or elsewhere). Some British recruits went to Australia to enlist after being rejected as a conscript by the British army. If isotopes had been utilised to assess provenance, then this would undoubtedly have led to false assumptions, as it would have categorised many of those enlisting in Australia as British, with the inherent assumption that they had fought for the British Army.

For the Fromelles ID process we had ante-mortem data (mainly from historical sources such as enlistment records) for most of the Missing (n=*c* 1650) and post-mortem data (data derived from their skeletons and associated artefacts) for the buried soldiers (n=250). The ante-mortem data varied enormously in quantity and quality and this has been touched on in previous chapters, in particular in relation to anthropology (see chapters one, four and six). Unfortunately, a large proportion of the military records relating to UK soldiers from the First World War were destroyed during the Blitz in the Second World War, while those for

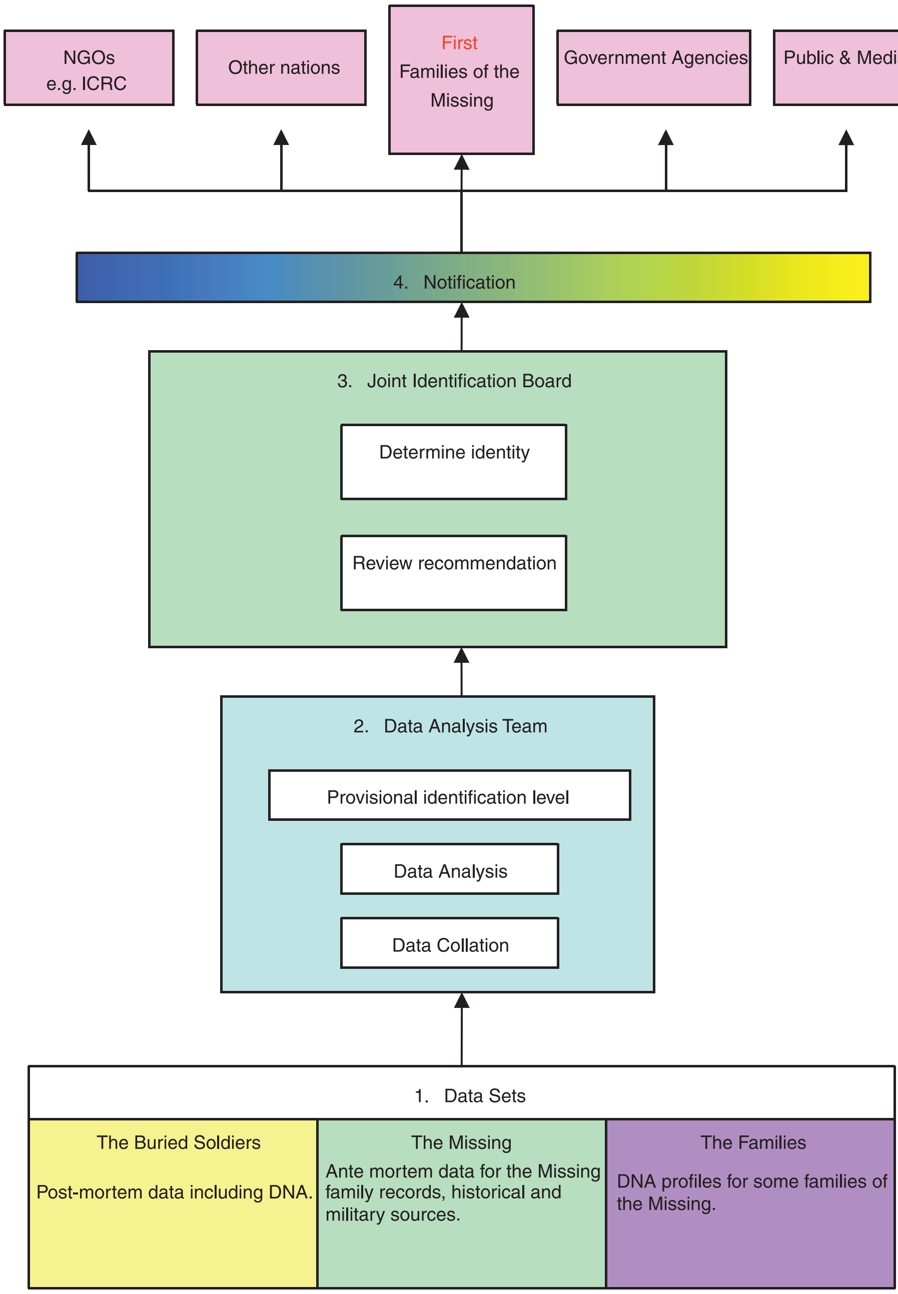

Fig. 7.1 Overview of the ID process used for the Fromelles Project from 2010 to 2014

the Australian army usually survive and can be extensive in quantity and character. This is a publicly available resource via the Australian National Library (http://recordsearch.naa.gov.au). For the British soldiers we do not even know the age at death for all of the missing soldiers, nor such characteristics as recorded stature, any information about health at enlistment or the cause of death. It is also important to appreciate that neither accurate medical records nor any detailed dental records have survived from this period. In fact, it is almost certain that such things were not routinely made at that time in any part of the world. Enlistment records, where they survive, can include some medical information, such as notes of scars or limps, comments about 'bad teeth' or the presence of 'false teeth'. Further, information about treatment received while in theatre was noted, but this was not detailed enough to be useful in most cases (for example, 'spent 3 weeks in the VD clinic', or 'had decayed teeth extracted'). However, in a few instances where it was recorded that a soldier had a serious wound or fracture to a specific part of his body in the weeks or months before his death, then this could potentially be very useful as this could then be compared against anthropological observations (see chapters one, four and six).

The value of anthropological observations in an ID process reflect that the biological parameters of human ID obtainable from the skeleton are a mixture of those that are inherited and those that are acquired during life and at the point of death. Those that are inherited include DNA, genetic disease, and handedness, and these are determined between conception and birth. Those that are acquired are the result of lifestyle, circumstances, environment, and life choices. They include such factors as diet, occupation, climate, living conditions, recreational activities, and exposure to interpersonal violence, injuries, disease, medical interventions, and of course fashion (see Chapter Four).

The greater the distance in time between the deaths of the Missing and attempts to determine who they are, the less relevant information survives, and in some cases, it may never have existed. Clearly where detailed and reliable medical and dental records, fingerprints, and family based information and photographs (ante-mortem data) survive for a missing person, this enhances the possibility of arriving at a positive ID.

The following issues have to be considered in any historical ID process, they must be weighted and then utilised on that basis:

1. In some cases, historical records and family information may be unreliable (for example, eye witness accounts of a soldier's death during battle; in some cases there was more than one for Fromelles and they did not tally; while family lore may be unreliable and distorted through time and retelling). With any historical source it is vital to first consider who wrote it, why they wrote it, when they wrote it and how did they know what they were writing about, i.e. were they witnesses or are they recording hearsay? Only after undertaking that process should one consider the veracity of the detail of what is written.
2. In the case of Fromelles, there was a paucity of ante-mortem records (particularly for those fighting for the British army) and importantly concerns about the accuracy of these where they do exist. We know, for example, that some men lied about their age at enlistment, and that this applied not only to younger individuals overstating their real age, but also to those who were over the specified upper age of enlistment who may have understated it.
3. For Fromelles, challenges and restrictions are imposed between the time of battle and burial. These include skeletonisation of the bodies and the sometimes fragmentary and the incomplete condition of the bone, plus an inherent potential degradation of molecular information.
4. The limitations of anthropological methodologies for determining such characteristics as age at death and stature must be appreciated. It is important to always consider which methods the anthropologists use, and whose version of that method, then to assess the confidence limits of each and the associated ranges and to work within the limits and constraints of those parameters.
5. Artefacts found with a soldier may not have originally been his. Transfer can occur for many reasons and the possibility and likelihood of this have to be considered (see chapters five and six).
6. It should be appreciated that although DNA is a very powerful tool, the DNA recovered from the buried soldiers at Fromelles was present in low quantities and in a degraded form as the result of poorly understood taphonomic processes. Fortunately, sufficient usable DNA survived to obtain Y-STR and mitochondrial DNA profiles for all the buried soldiers.
7. Unfortunately, informative living family members (those sharing the same paternal or maternal DNA as the buried soldier) do not always exist today, or may not be able to be traced in all cases, and in some cases where they are they may not be willing to donate samples. A very small number of people have declined to participate thus far. Of the British families asked to donate, three of 277 (0.9%) refused.

For Fromelles, challenges imposed by all of the above were exacerbated by the fact that we have 250 skeletons (buried soldiers) and approximately 1650 missing soldiers. Therefore, we potentially have 6.6 times more ante-mortem datasets (of varying amounts and reliability) than we do post-mortem datasets (of varying amounts and degrees of confidence).

Clearly, biological evidence as derived from the buried soldier's skeleton is more reliable and is more heavily weighted in the ID process than any artefactual evidence found with him. This reflects the fact that inscribed items such as 'dog tags' or personal effects that may be found with a skeleton may not necessarily have been the property of the buried soldier with whom they were recovered. As discussed in Chapter Five, we consider many such objects to be portable, and transfer from one to another person may take place for a variety of reasons (theft, lost in a bet, picked up on the battlefield, gifted, swapped). An example showing the complexity of this was in one case where a tobacco pipe was found with a buried soldier (1859B; see Chapter Five), with one set of inscribed initials over-inscribed by those of another, as revealed during further analysis undertaken during the DAT; the ID process demonstrated that this second set of initials matched the name of the soldier with whom the pipe was found (see Fig. 5.39). It was for that kind of reason that the archaeologists placed emphasis on the location of items, in addition to the assemblage of items found with individuals, rather than on the ID significance score assigned to each item.

The British Army and its Commonwealth forces issued ID (dog) tags from the beginning of the First World War. The tags were stamped with the name, rank, and regiment of the soldier and were initially made of fibre and suspended around the neck by butcher's twine. The army changed its regulations on July 6, 1916, so that all soldiers were issued two tags, one to stay with the body and the other to go to the Red Cross in the event of death in battle. Despite this, the soldiers at Pheasant Wood mostly appear to have only been issued with one dog tag, and after the battle, the Germans diligently removed these from the dead on the battlefield, along with any other personal belongings they could find, and returned them to the Red Cross in order that records could be made and relevant families informed. It is fortunate that some of the buried soldiers had seen fit to have a second dog tag made and that three of these survived (see Figs. 5.35 and 5.36), though not all were complete or legible. Very few inscribed artefacts recovered were considered to be unlikely to be portable, but those that did included bespoke (that is, not off the shelf) false teeth, one set of which was inscribed with the wearer's name (Fig. 7.2 and see Chapter Four).

As noted above, determination of a Category 2 level ID (the soldier is identified to the army for which he fought) is likely to reflect artefactual evidence, such as army or regimental items. This was considered in the light of contextual data (that is, where on the body the item was located) to ensure as far as is possible, that the item was a part of the buried soldier's uniform. Hence, items found in the region of pockets, which may or may not be with the original owner, may be discounted as not being a strong enough indicator. The most useful artefact of all in this process was the jacket belt buckle, which was sewn onto the belt at the time of manufacture, which in turn was sewn into the seams of the jacket. Sixty-five per cent of all of our buried soldiers had Australian jacket belt buckles (see Chapter Five).

DNA and ID

Alex Jeffrey developed DNA profiling in 1984. His discovery was that particular detectable DNA segments show variation and followed the rules of inheritance discovered by Gregor Mendel in the 19th century. In the late 19th and early 20th century genetics was in its infancy. Friedrich Miescher discovered DNA in 1869, and Thomas Hunt Morgan demonstrated the inheritance of genes in chromosomes. However, the discovery of the structure of DNA had to wait a further 40 years for the work of Watson, Crick, Wilkins and Franklin. DNA has some unique properties, particularly the way it is inherited from one generation to the next. Each cell within the body contains a nucleus, within which are the chromosomes, which act both as a blueprint and as a structure, enabling cells to divide and propagate. There are 22 pairs of chromosomes in each cell, plus the sex chromosomes. They consist of a pair of X chromosomes in females, and one X chromosome and a smaller Y chromosome in males. A DNA profile records the similarity or difference in the DNA sequence between each individual as single base pair changes in the case of the mitochondrial sequence, or the number of short tandem repeats (Y-STR) in the case of the Y profile. It must be stressed that DNA is not a 'golden bullet' for determining ID in cases of unidentified people who died almost a century ago. However, it proved to be very useful as the initiation point for the identification process. This project was the single largest undertaking of its type to date for cases of this era, but it was only with the DNA profiles of living relatives of the missing soldiers that an assignment of ID could be made for the buried soldiers. In some specific cases it was also possible to go back to samples from deceased family members (with consent and support of the family), and recover DNA from medical slides and biopsies. In line with our process, it must be stressed that an ID indicated by a high match probability is only confirmed after all the evidence from the historical, archaeological and anthropological record has been considered and is supportive and not contradictory. To date a high match probability obtained from a match on the maternal and paternal side has not been contradicted, but there have been many cases where a single match with low match probabilities was contradicted by other forms of evidence.

Fig 7.2: Denture showing negative impression of the wearer's name

Establishing inheritance patterns between an individual and relatives separated by almost a hundred years, and two or more generations is a challenging task. The chromosomal genetic markers that form part of the National DNA database and are used to link a person to a scene of crime were not suitable for the Fromelles project as they are diluted and recombined through the generations, essentially rendering them of no practical use in ID over extended generations. This problem is overcome by looking at the inheritance pattern of those markers that are passed from one generation to the next and remain essentially unchanged or 'un-shuffled'. These markers are those on the Y chromosome (which determine the male lineage), and the variation in mitochondrial DNA sequence (which determine female lineage). The DNA profile can thus be followed through the generations and lineage established. This principle works very well for mitochondrial DNA as this does not have a high mutation rate but can be more problematic for the Y profile as this can mutate over a small number of generations (Hohoff *et al.* 2007). Even so, some profiles with one mismatch can have very high confidence limits due to their uniqueness and can be a very strong indicator of relatedness. By this mechanism the expected DNA profiles of the missing soldiers was established (assuming non-adoption and legitimacy).

The strongest match probabilities that we have so far obtained for our buried soldiers, *vis-à-vis* those of the missing soldier, are 1:1915 for the mDNA and up to 1:3387 for the Y-STR. Taken in isolation such confidence limits are insufficient to yield an ID at the level of 'beyond reasonable doubt', or as prescribed for our category 1. Where we have both a paternal and maternal match for a buried soldier then the two statistical vales are multiplied and this can yield match probabilities of up to 1:6.4 million, which is more acceptable. However, in many cases we might only have a donor from one line or the other and the match probabilities are very low, typically <1:500. This is typical of situations in which the family had a particularly common mitochondrial profile. In such cases, we rely upon the project historians and family liaison colleagues to attempt to trace an informative donor on the contra line; in some cases, this has been successful.

In order to establish a familial link it is essential to locate living genetic relatives from, if possible, both their maternal and paternal lines. Without living relatives it is extremely difficult to match the DNA of a buried soldier to that of the family of a missing soldier, as the buried soldier's DNA profile, without an external reference, is simply an anonymous signature. In practice, this can be a complicated process as many of the missing soldiers were young unmarried men, with no (known) offspring. In such cases appropriate genetic relatives have to be traced from either male descendants of his brother(s) or female descendants of his sister(s), or, if he had none, or they had none, from descendants of his father's brothers' male line or his mother's sisters' female line. When selecting individuals from families to provide a DNA sample, only those 'informative' individuals on either the direct male or the direct female line are asked to be DNA donors. Some descendants of the soldier will not have useful or 'informative' DNA in relation to that of the missing soldier, as they are not direct male or direct female descendants, despite being close family relatives. As the DNA markers are essentially unchanged from one generation to another, it is possible to include very distant relations to help establish a common ancestor. This will allow identification of a number of different branches of the family and then the line can be followed forward to the current generation. There are some practical limitations with this technique but it is nevertheless very important in tracing distant relatives if no known informative family members are still alive (see Figure 7.3). Often the individual who has been traced by the genealogist is unaware of their relationship to the buried soldier.

Tracing families

The Joint Casualty and Compassionate Centre (JCCC) in the UK, and the Unrecovered War Casualties section of the Australian Army undertook tracing and contacting living relatives (see Chapter One). In Australia, the Fromelles Project has had huge amounts of publicity reflecting its importance in Australian history, and many people came forward as a result. This battle represented the greatest loss of life the country has ever witnessed in a single event and the fact that so many soldiers remained missing at the turn of the new millennium was a matter of great concern to some (Steel 2010), their efforts ultimately leading to this project. In the UK, the Battle of Fromelles was almost unheard of until 2009 and as such a considerable amount of effort was necessary to make people aware of the project and actively seek families and descendants of the missing British soldiers from the battle. Regimental groups and societies were most helpful and the project is particularly indebted to the efforts of Mel Pack who undertook a great deal of expert genealogical research, tracing most of the British families and putting them in touch with JCCC. The media, through articles about individual soldiers in regional papers, has been particularly successful at generating interest, and living relatives have been identified via that route.

Road shows were held in specific regions of the UK explaining the nature and aims of the project and much interaction with the public and media took place. All those considering donating DNA were made aware of the aims and objectives of the project and issues of informed consent were discussed and agreed. Great care was taken throughout to manage expectations, so that those donating a sample were aware of the relatively low possibility (1:6.6) of their long lost relative being discovered and so would not be too disappointed if no positive match was made. Apart from asking families of the missing soldiers for DNA samples, we also issued UK soldiers' families with a questionnaire to attempt to glean further information to supplement that existing in surviving records. While we appreciated that this was unlikely to yield very much information, one fact could, in theory, lead to an ID so it was worth the effort. Further, some families have good photographs, which they kindly copied and sent to us as well as occasionally being able to suggest something useful in terms of, for example, the soldier's occupation before enlistment. Unfortunately, many photographs of missing soldiers show them with their army cap on, covering hair and casting a shadow over their eyes; head shape is also not clear in such cases. The most useful photographs to set against the 360º videos (chapters two and four) are those where the face was at an angle, showing the facial profile more clearly than a full-face image. Where this was most useful was in the case of individuals with unusual facial proportions or shapes, particularly of the chin and nose.

DNA haplogroups

DNA also serves another useful purpose in ID as it holds clues to our distant ancestry and, where this might be unusual, it can provide useful data on ethnicity, but not on identity. This reflects that Homo sapiens evolved around 200-150,000 years ago in East Africa, and the species has steadily expanded to occupy the entire planet. Given the limitations on travel historically, there were limited opportunities to form relations with distant populations, and consequently particular DNA profiles – haplotypes – are common amongst certain ethnic groups or in particular geographic regions. Grouped together these form a haplogroup. From an individual's DNA profile, their haplogroup is obtained and this will indicate a particular geographical origin or ethnic group. The norm for individuals from an immigrant Australian population at the turn of the 20th century would be a north-western European haplogroup and where common haplogroups are indicated, they are of little value in ID. In this project, however, the ability to determine if someone had Aboriginal, Jewish, Eastern European or Scandinavian ancestors within the family most certainly is. Jewish ancestry is important because religion was noted in enlistment records. Aboriginal ancestry was helpful as while ancestry was not recorded at enlistment, some Australian recruits were noted as being 'dark' or being known as 'darky' (see Chapter One) and this term was commonly used at the time to suggest Aboriginal ancestry. Surnames indicating, for example, those from Eastern European or of Scandinavian ancestry could also be compared against those with such haplotypes. It must be stressed, however, that this technique is not reliable, nor is it used for assigning ID; it only suggests a particular ancestry and might support an ID. Neither does it determine how far back in time a particular genotype came into a family, or indeed how it got there, or under what circumstances.

As an example of this, the first author of this chapter has a maternal haplogroup of Xb2. This haplotype diverged about 30,000 years ago and is unusual in the UK. It has been found most commonly among the indigenous First Nation population of the Algonquin region of North East America, and there is also an association of the same haplogroup with central Europe and the Middle East. As far as she is aware, her recent maternal ancestry (back into the 18th century) is Southern Irish so this must have been her maternal line far too long ago to be helpful in determining her ethnicity; in fact, it could be very misleading.

Anthropology in identification

Anthropological analysis is very important in human identification as it can provide an assessment of such critical facts as age at death, ethnicity, living stature and, of course, it can provide evidence of some chronic disease processes or healed trauma, all of which may be noted in enlistment, in medical or other records (see also Scheuer and Black 2007). Key parameters for this project are age at death and stature estimation, as while the latter is recorded at enlistment, the former can be deduced from such records *vis-à-vis* date of death. That said, both methods have limitations in their use.

The first, and common to both, is the issue of accuracy of recording when the written record was made. As noted above, it is known that the very young, and those older than the enlistment age could and did lie about their ages, and at a time when no birth certification was required at enlistment, where the man looked older or younger than he really was no-one other than he would appreciate the inaccuracy. It is also well attested that at certain times when more soldiers were needed by the armed forces, a blind eye would be turned to obvious falsehoods. Stature too is something that can be inaccurately measured and/or recorded; further, living stature is greater in the morning than in the evening, so time of measurement is crucial.

The other great limitation in the weight to be attributed to anthropological data is the lack of accuracy in the methodologies, particularly for those dying over the age of thirty. Over the last 30 years, several projects have taken place where the 'real' age of individuals was known (Saunders *et al.* 1992; Molleson and Cox 1993; Loth 1995; Cox 2000; Mulhern and Jones 2005). These were then examined in relation to estimated ages and the results clearly showed that while such methods are useful for indicating broad trends, and very useful for population studies, they are far less helpful for individuals once physical maturity has been reached. For age at death, the lack of the ability to provide a meaningful age range for identification purposes is far greater at certain ages than at others. For example, the range of possible ages in those dying up to the early twenties, and around thirty, is much tighter than for those over 35 where the age range suggested by several methods is many decades (and then only at 95% confidence limits, that is, 1 in 20 will fall outside that range). Further to this, where several ageing methods are utilised for individuals, each suggests a different mean age and offer different ranges, hence the age range proffered by the anthropologist is averaged out of the suggested ranges and no precise method exists for achieving this. While some anthropologists consider that a multifactorial approach helps to minimise errors (Jackes 1992), others argue that it may compound them (Cox 2000). When buried soldiers are recommended for a Category I ID, while the real age of the majority is usually somewhere within the suggested range, for others they could fall at or just outside one end or the other of what is a very large age range. Age at death estimation is just that, an estimate, and for younger adults it is usually accurate to within a few years, being much less precise for older men.

Stature estimation has limitations as mentioned above (see Simmons and Haglund 2005), but was generally very useful in this project, with most soldiers' stature falling within or very close to the plus or minus 3-4cm range offered by the methods used. Where estimated stature diverged from recorded stature ranges, then scrutiny of the anthropological record was made to see if there was any reason that might explain why this should be so. Occasionally, divergence in corresponding limb length would be noted or such factors as shortening of a limb due to a badly healed fracture. The inherent impact of such wide ranges of possible age and/or stature estimates, as examples of the imprecision of results derived from anthropological analysis, are summed up by Byers (2008, 16): 'One of the most common problems faced by forensic anthropologists is how to make a single determination from ambiguous data.'

Examples of ante-mortem trauma and pathology identified among the soldiers are described in Chapter Four. Some of these, for example cribra orbitalia, were of no real use in the identification process, because they are asymptomatic conditions in themselves, and in any event are not recorded on ante-mortem records, as many probably went unnoticed in life (see chapters four and six). However, these drawbacks considered, there were occasions when ante-mortem trauma and pathology provided good supporting evidence. For example, a limb abnormality was observed on one skeleton and was recorded in the ante-mortem records for the same individual. In addition, some of the dental treatments and/or conditions correlated with enlistment records where gold fillings, dentures, caries, 'bad teeth' and/or previous rejection on account of poor dental health were noted. There was also a higher than expected (considering inaccuracies in recounting events – see above and chapters four and six) correspondence between some eye-witness testimonies and the peri-mortem trauma observed on the skeletons. This was identified at the assemblage level (see Chapter Six), but also on an individual basis during the DAT. While ante-mortem trauma and pathology were by no means key forms of identification evidence, they could provide important supporting evidence in some cases.

Identification assisted by the program OPTIMISE

Our analysis was helped greatly by the use of OPTIMISE, a statistical program based on nearest neighbour analysis, developed for this project by Richard Wright. This program produces optimised spreadsheets, and to assist in identification, it uses two variables, age at death and stature. Age and stature are of course on different scales, so before calculating distances, OPTIMISE standardises these values. This process gives equal weight to both variables in the computation of distance.

Age and stature are independently recorded in two datasets:

1. For 248 of the 250 buried soldiers there are anthropological estimates of age and stature.
2. Age and stature are included in the service records for 1268 AIF and British missing soldiers. Each individual has a line in a spreadsheet that gives additional information that may leave osteological markers, such as occupation, ante-mortem injuries, and cause of death.

The purpose of OPTIMISE was to take anthropological estimates of age and stature for each of the 248 buried soldiers with available data. The process then searched the 1268 rows of the service records for those individuals that are closest to the excavated individual in terms of both age and stature; this was repeated for each buried soldier. The usefulness of this model is illustrated in Figure 7.3, which shows the age at death and stature taken from the service records of the 1258 AIF soldiers thought to be possible targets for identification. The red asterisk represents the age and stature of one of the 250 buried soldiers. The nearest neighbours from the service records, in terms of the joint consideration of age and stature, are included in an arbitrarily drawn red circle. For the identification process, desirable information is the names, from the service records, of those nearest neighbours within the red circle. If we know these names, then for the purpose of identification we preferred to look first at their properties. By contrast, it would be inefficient to examine the distant individuals that lie, for example, at the bottom left corner of the distribution of the 1,258 soldiers represented in Figure 7.4.

The program OPTIMISE was more thorough in its approach than merely providing the names of nearest neighbours within a small circle. OPTIMISE uses what in statistical terms is called nearest neighbour analysis. This process considered the joint space of age and stature. The direct distance was calculated from the buried soldier's age and stature to each of the 1268 individuals in the service records (there are 1258 AIF and 10 British soldiers). OPTIMISE then

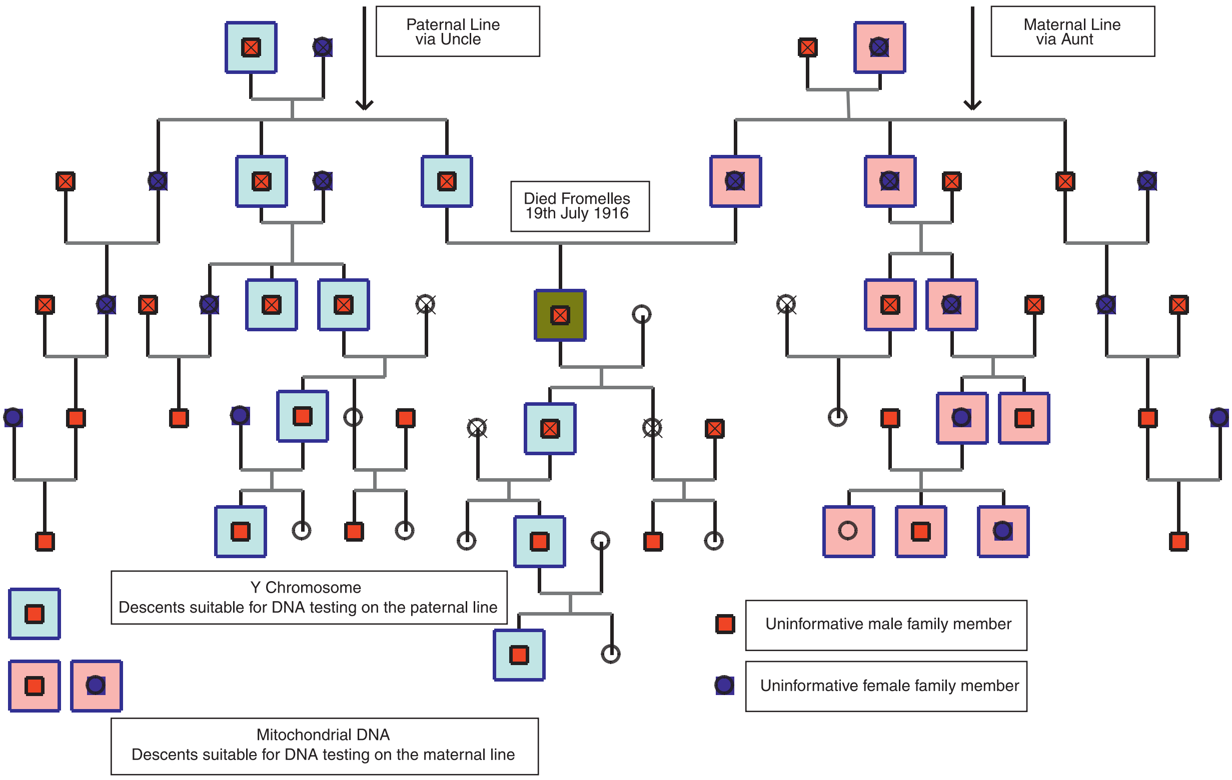

Fig. 7.3 Pattern of inheritance of the Y-STR and Mitochondrial DNA showing potential paternal and maternal donors. Those on the direct male line are shown as blue boxes, while those on the maternal line are shown as pink boxes

used these distances to sort the 1268 rows of service records, so that they were now ordered at increasing distance from one of the 248 soldiers. By this process we obtained unique spreadsheets for each of the 248. In each spreadsheet the first row is the nearest neighbour of a particular soldier and the last row (row 1,268) the furthest neighbour.

Although our approach to identification began ideally with a strong DNA match, where none existed then OPTIMISE helped by highlighting or optimising individuals in the top few rows of the spreadsheet as possible candidates for a particular missing soldier. It was also very useful where multiple matches to individuals with relatively common DNA matches occurred, as it helped rule individuals in or out relatively easily. The stopping point for examining the rows was normally where ages and statures, as documented in the service records, were judged to be unreasonably divergent from the estimates made by the anthropologists for the single individual. The only time this was deviated from was when a disease or other skeletal characteristic might have led to one of the buried soldiers not conforming to the normal characteristics for age and/or stature estimation. It is clear from Figure 7.4 that this approach was most efficient at assisting identification at the fringes of the distribution of both age and stature.

There are of course confidence limits to consider in the anthropological estimates of age and stature and there may also be errors in the service records. These uncertainties create noise, but do not affect the efficiencies of time that OPTIMISE offers when undertaking the identification process. In summary, OPTIMISE did not make identifications. Rather it offered the potential to speed up the process by utilising age and stature to link the remains of the buried soldiers with those missing soldiers with service records.

CONCLUSIONS AND OUTSTANDING QUESTIONS

At the time of writing, the JIB has confirmed the identification of 144 of the buried soldiers from the list of over 1650 missing soldiers. One hundred and six of the buried soldiers remain without a name and of course many hundreds are still 'missing', their place of burial unknown. Of the 106 still not identified to a name, 75 are considered to have fought for the Australian Imperial Force, two for the British army and 29 remain 'known unto God'. Each year, on 19th July, any new IDs are formally named with an inscribed headstone and full military honours at the new cemetery in the village of Fromelles. This process, attempting to determine the identification of the 106 will continue but not under the auspices of the FMB. Going forward, subsequent ID analysis processes have reverted to being a national responsibility with each country assuming responsibility for its own missing soldiers.

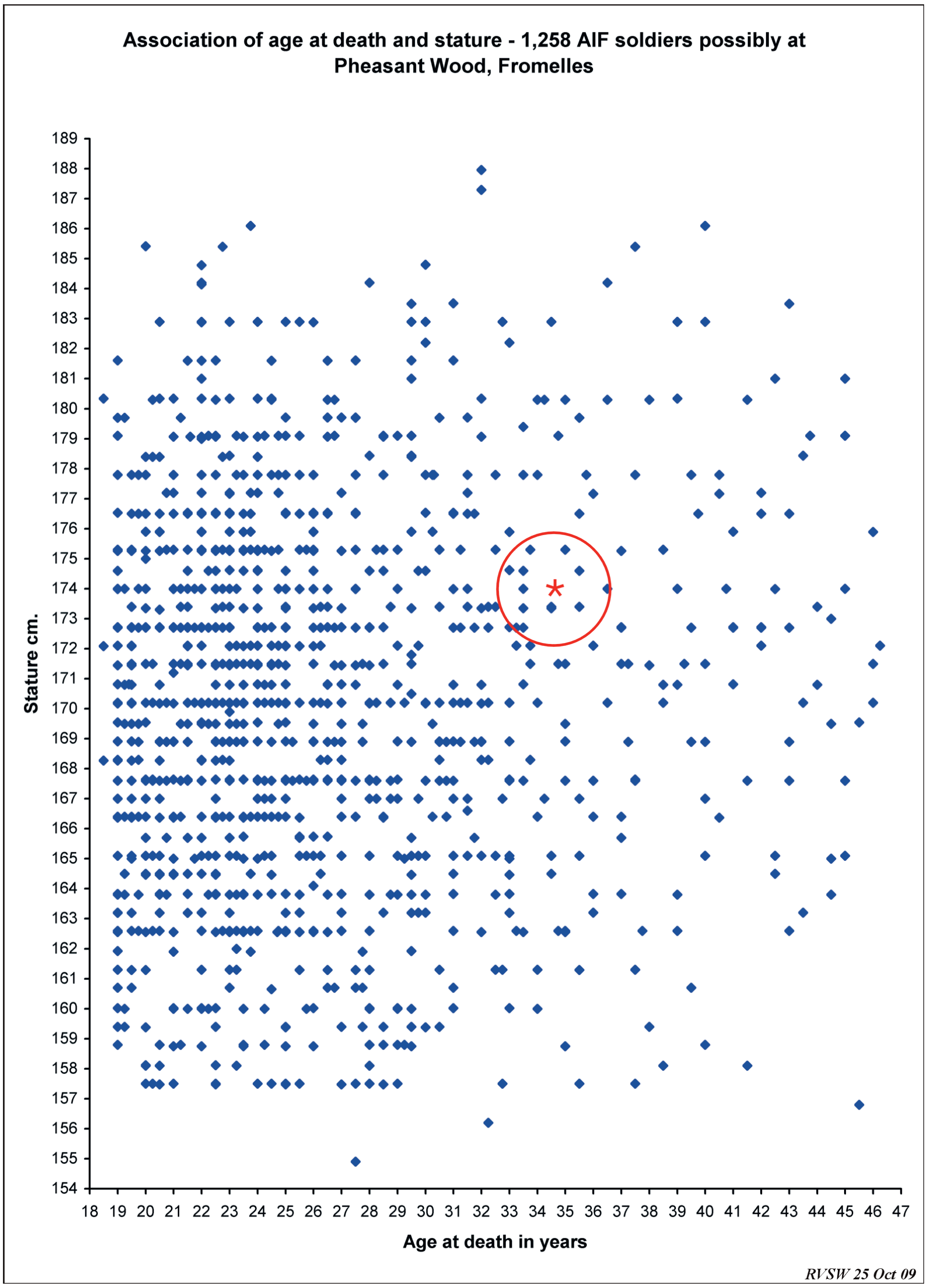

Fig. 7.4 A visual representation of the program OPTIMISE

The lack of identified British soldiers fighting with the British 61st Division of the British Expeditionary Force and hailing from Gloucestershire, Worcestershire, Warwickshire, Berkshire and Buckinghamshire is puzzling. Some historical accounts, and some historians, suggest that they were buried with those from the Australian 5th Division in the mass graves dug by the German burial parties adjacent to Pheasant Wood (Steel 2010). A document housed in the Bavarian military archives and written by Major-General von Braun, commanding RIR21, specifically ordered the construction of mass graves for up to 400 English soldiers, though a Red Cross document only mentions the burial of 160 Australian soldiers near Pheasant Wood and a Red Cross record from the file of Lt John 'Jack' Charles Bowden (59th Battalion) specifically mentions five mass graves near Pheasant Wood, Fromelles. The lack of British soldiers identified to date may be explained in a number of ways. The first is that the order of von Braun was never carried out and that any British dead recovered by the Germans were buried elsewhere; they were fighting to the west of the Australians so this is not impossible. Secondly, the lack of named British soldiers may be an artefact of our ID process. A DNA match to a British soldier, with weak match probability, combined with the lack of precise historical records (consequently useful ante-mortem data), makes us almost totally reliant upon DNA and we did not have strong enough matches with which to recommend an identification. Where DNA results suggested a British soldier but with a low-match probability then we needed something convincing and relatively unique in order to add weight to that result before we could feel confident in recommending a British ID. Thirdly, we may not have had the families and donors for the particular British soldiers in our cohort, or may have only had one side of the family, which has resulted in no matching profile, or at best a weak and inconclusive DNA result.