Managing forensic DNA records in a divided world: the Belgian case

Patrick P.J.M.H. Jeuniaux, Bertrand Renard, Leen Duboccage, Séverine Steuve, Caroline Stappers, Inès Gallala, Sabine De Moor, Alexia Jonckheere, Benjamin Mine, Beatrijs Vanhooydonck, Morgane Kempenaers, Christine De Greef, Pierre Van Renterghem and Vanessa Vanvooren

(Author affiliations can be found at the end of the article)

Abstract

Purpose – This paper aims to describe the activity of managing records related to forensic DNA identification. First, it illustrates the fundamentals behind the technique of forensic DNA identification. Second, it explains the legal and institutional contexts in which it is used as well as the notion of DNA-based judicial records. Third, it provides details of records management issues that are met in practice.

Design/methodology/approach – An interdisciplinary team reflects upon the practices surrounding the management of DNA-based records in the Belgian National DNA database during more than 10 years.

Findings – The main problems with managing DNA-based judicial records stem from the existence of natural boundaries between the various stakeholders operating with or within the Belgian judicial system. Six types of issues have been found: non-automaticity and omission, error-prone and inefficient manual operations, electronic issues, results quality, useful reporting and incoherence and duplication. These problems are discussed in terms of four records characteristics: completeness, correctness, traceability and usability.

Research limitations/implications – The research is limited to the Belgian case with no comparison with other countries.

Practical implications – This paper attempts to formulate general principles that aim to stimulate good practices in managing records in the field of criminal justice.

Social implications – The ethical issues surrounding the domain of criminal policy (e.g. the proper use of financial resources, the fair and balance use of records to carry out justice) are of general interest to the public.

Originality/value – The paper benefits from a large temporal angle (more than 10 years) and applies a multidisciplinary viewpoint on its subject.

Keywords Criminal justice, Records management, Public policies, Administrative boundaries, DNA database, DNA-based judicial records

Paper type Research paper
Introduction

In this paper, we describe the activity of managing the criminal judicial records that are related to the technique of forensic DNA identification (hereafter, the “DNA-based judicial records”). To our knowledge, such a description is absent from the literature in records management. Given the societal worries focusing on the use and misuse of technologies involving personal data such as those based on DNA, whether they are concerns for the right to privacy (Renard, 2011; Rothstein and Talbott, 2006; Shapiro and Weinberg, 1990; Williams and Johnson, 2005), function creep (Dahl and Sætnan, 2009; Renard, 2008), inefficiency of justice (McCartney, 2005), abuse of law-enforcement or miscarriage of justice (Machado et al., 2011), a thorough inspection of the mechanisms that handle DNA-based judicial records should be of great interest to a democratic society (Renard, 2013; Williams and Wienroth, 2014). For instance, issues of records retention may induce infringement on private life as was seen in the judgment of the European Court of Human Rights “S & Marper vs United Kingdom” (cf. Prainsack and Toom, 2013). By opening up a rare window on the processes that handle such records, we aim to offer a contribution to a much needed collective and public debate.

Taking Belgium as the context, we explain that DNA-based judicial records used in that country are composed of different elements that are spread across several administrative boundaries or stakeholders and that this divided world creates specific issues of records management. We show that the ability of correctly linking the different elements that compose DNA-based judicial records is a key objective of proper records management. We also explain what processes attempt to ensure that records are complete, correct, traceable and usable.

Methodologically speaking, our description of the issues is set up in the context of the Belgian National Forensic DNA Database (hereafter “the national DNA database” or “NDD”) within the arena of the criminal justice system. Among the authors of this paper, some have worked at the NDD department from its onset, while others are practitioners currently operating within it (records managers, database experts and biochemists) – spanning together more than 10 years of activities. To reflect on their experience, they teamed up with colleagues (data scientists, criminologists) who investigate topics relevant either to the functioning of justice and society at large or the use of forensic evidence in particular. Parts of the thoughts and observations that are shared in this paper have been collected either from the literature or from participant observation and interviews of relevant stakeholders which have been previously reported (Renard, 2008).

Forensic DNA identification

Forensic science can be defined as the exploitation of science to support law-enforcement and judicial decisions or as the study of crime and its traces (Roux et al., 2012). Since the discovery by Jeffreys et al. (1985) that individuals can be uniquely characterized by their DNA “fingerprints”, forensic DNA identification is used to support criminal justice. A variety of technologies has been developed since then. The standard technique used today leads to characterize individuals in terms of their “DNA profile” (Gill and Buckleton, 2004).

DNA profiles

For illustrative purposes, fictitious examples of DNA profiles are given in Table I. The headers of the table from L1 to L12 represent 12 “loci” (i.e. DNA is an elongated molecule...
In which specific locations called “loci” can be characterized; real loci would bear complicated names such as “D3S1358” or “vWA” instead of “L1” or “L2”), whereas the last column represents the gender of the person who is at the origin of the analysed sample. DNA profiles are obtained by analysing biological materials (e.g. saliva, blood, hair, skin). In Table I, the three rows underneath correspond to profiles that were established on the basis of three physically distinct biological samples (for example, a saliva swab, a sample of blood and a tuft of hair). The goal of a forensic DNA expertise is to determine if some of these biological materials could originate from a same person. Each cell in Table I contains a pair of numbers (the two “alleles”) that characterizes a particular property of DNA at that particular locus. We have two numbers because one relates to the mother and the other one to the father of the individual.

Profiles that are established on the basis of stains found on crime scenes are called “stain profiles”, whereas those that are established by analysing samples taken directly from known, referenced individuals are often called “reference profiles”. The most immediate utility of forensic DNA identification is that it allows linking a stain profile to a reference profile (or showing that there is no link, e.g. excluding suspects from consideration). Using this technique, it is therefore possible to identify the putative donor of a stain if a reference profile is available.

When the profiles are different, it is concluded that they belong to different persons. In the examples of Table I, we see that Reference Profile 1 (Ref1) is different from Ref2 and therefore they cannot belong to the same person (assuming there are no errors in the profiles). Although Stain Profile 1 (Stain1) has a smaller number of analysed loci than Ref1 and Ref2, it does not prevent their comparison, for a comparison of two profiles is based on the loci they have in common (here, loci 1-8). Ref1 and Stain1 match on all the values they have in common. Ref2 and Stain1 have the same values for Loci 1-7, but they do not match on L8; hence, they cannot originate from the same person.

When profiles match, a statistical reasoning is conducted to determine how much more likely it is to obtain such a result if the profiles come from the same individual than if they come from two different persons (Aitken et al., 2010). As Ref1 and Stain1 have identical values for the loci they have in common (Loci 1-8), they may originate from the same person, but they could also originate from two different persons who, by coincidence, share identical values for the analysed loci.

We therefore see that there exists a possibility to obtain a positive result (i.e. a match) in favour of the hypothesis that the two profiles came from the same person, whereas the truth is that they came from two distinct persons. This is called a false-positive. The other type of error is the false-negative, where we fail to find a positive result in the case where two profiles actually came from the same person.

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>L1</th>
<th>L2</th>
<th>L3</th>
<th>L4</th>
<th>L5</th>
<th>L6</th>
<th>L7</th>
<th>L8</th>
<th>L9</th>
<th>L10</th>
<th>L11</th>
<th>L12</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ref1</td>
<td>1,21</td>
<td>8,9</td>
<td>1,7</td>
<td>6,7</td>
<td>1,3</td>
<td>7,9</td>
<td>5,6</td>
<td>2,11</td>
<td>7,19</td>
<td>4,13</td>
<td>1,8</td>
<td>7,24</td>
<td>Male</td>
</tr>
<tr>
<td>Ref2</td>
<td>1,21</td>
<td>8,9</td>
<td>1,7</td>
<td>6,7</td>
<td>1,3</td>
<td>7,9</td>
<td>5,6</td>
<td>7,8</td>
<td>7,10</td>
<td>4,13</td>
<td>7,9</td>
<td>1,23</td>
<td>Male</td>
</tr>
<tr>
<td>Stain1</td>
<td>1,21</td>
<td>8,9</td>
<td>1,7</td>
<td>6,7</td>
<td>1,3</td>
<td>7,9</td>
<td>5,6</td>
<td>2,11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Male</td>
</tr>
</tbody>
</table>

Note: Columns represent loci, rows correspond to profiles and cells contain alleles.

Table I. Fictitious examples of forensic DNA profiles

Managing forensic DNA records
Databases

In the 1990s, the technology of DNA databases was introduced and rapidly spread thanks to scientific, computer-related, normative and legislative evolutions (Walsh and Buckleton, 2004). Forensic DNA databases keep records of DNA profiles for criminal justice purposes, and allow making comparisons between profiles that do not belong to the same criminal affair, therefore extending the possibilities to discover links between two affairs or between an affair and a reference profile. Moreover, it is a resource that may be exploited for broader intelligence seeking and criminological research purposes (Jeuniaux et al., In Press; Lammers and Bernasco, 2013; Lammers, 2014; Ribaux et al., 2010; Rossy et al., 2013).

The usefulness of the data recorded in forensic DNA databases depends on an elaborated chain of human actions (e.g. emitting a judicial decision, sampling the biological material, producing DNA profiles, transmitting, storing, maintaining, exploiting and deleting the records) that is shaped by scientific, organizational and legal constraints (Gill and Buckleton, 2004; Ribaux et al., 2010). The specific way this chain is shaped in the Belgian context as it changes over time and what it means in terms of records management is the topic of the following sections.

Legal and organizational contexts

Soon after the discovery of technology in 1985, the exploitation of forensic DNA began in Belgium, opening the gate to an area of activity regulated by laws not specifically designed for handling this type of evidence. This original situation led to a diversity of practices, a relatively large number of laboratories producing DNA profiles, and the creation in some of those laboratories of convenient local DNA databases holding DNA profiles established for the judiciary, as well as ad hoc information relevant for supporting criminal investigations. The laboratory experts could spontaneously compare newly established profiles to those already stored in their local database, leading to important clues for identification purposes or discovering that a same person was involved in distinct criminal affairs by linking the DNA profiles that belong to that person (Renard, 2008; Renard et al., 2013).

Following a decade during which laboratories invested in forensic DNA techniques and introduced them in support of criminal justice, the Belgian legislator provided a specific legal framework to regulate their use (Moniteur belge, 1999, 2002). We refer to this legislative innovation as “Framework 1” (for a discussion, see Renard et al., 2000). Before Framework 1 took place, a laboratory expert would not contact experts from other laboratories to proceed to inter-laboratory comparisons of DNA profiles unless requested by the judiciary – a situation which limited the ability of magistrates to discover links between criminal affairs. To systematize the discovery of such links, and better ensure data protection and privacy concerns, Framework 1 triggered the creation of a national DNA database (NDD) to centrally store and regulate the use of the DNA profiles established by a group of nine authorized DNA laboratories (Jeuniaux et al., In Press; Renard et al., 2013). Contrary to countries where forensic DNA is under the direct control of law-enforcement, the Belgian legal framework has been designed to guarantee its control by the judiciary. DNA profiles are established by an authorized DNA laboratory and transferred to the NDD only by explicit request formulated in writing by a magistrate. The judiciary has the authority to request services from the NDD department (e.g. the comparison of a profile from a specific judicial case to the profiles
stored in the NDD). This department is located in the national forensic and criminology institute (Nationaal Instituut voor Criminalistiek en Criminologie/Institut National de Criminalistique et de Criminologie – NICC/INCC).

By law, a variety of administrative information must be associated with the DNA profiles stored in the NDD, among which the sample ID, criminal affair ID, person ID and any link between this profile and other profiles in the NDD (Table II). Neither the names of the identified individuals nor details about the criminal cases are available in the NDD.

What we call a “DNA-based judicial record” is a composite entity whose elements are created or stored by at least four types of stakeholders:

1. the judiciary (which holds information on criminal affairs and identification of persons, and which is organized according to 27 geographically distinct judicial districts, a federal district and 5 Courts of Appeal, i.e. 33 authorities in total);
2. the police forces;
3. the DNA laboratories; and
4. the NDD department (see Table III for a schematic idea of the whole records management process as it may ideally unfold).

Other relevant stakeholders would include court registries (that store forensic evidence), the designers of forensic toolkits, etc.

All of these stakeholders are to be understood as divisions or boundaries across which information must be thoroughly and efficiently exchanged. This exchange of data is needed because no stakeholder holds all the information. The distributed character of the information serves several purposes:

• it organizes work according to professional specialities;
• stimulates the independence of the stakeholders;
• imposes a limit on their power; and
• confers a neutrality and objectivity to the forensic and evidential process.

However, this state of affairs also generates specific issues that we discuss in the section “Records management issues”.

<table>
<thead>
<tr>
<th>Identification no.</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Person ID</td>
<td>Identification number referring to a person whose DNA profile has been analysed and is considered as a “reference”. It allows a magistrate to retrieve the identity (i.e., name) of the person</td>
</tr>
<tr>
<td>Sample ID</td>
<td>Identification number of a physical sample analysed by a DNA laboratory. The analysis of that sample may lead to establishing a DNA profile or not</td>
</tr>
<tr>
<td>Criminal affair ID</td>
<td>Depending on the nature of the affair, one or more references to criminal cases in the area of the justice department external to the NDD</td>
</tr>
<tr>
<td>NDD case ID</td>
<td>Identification number of the DNA judicial records under which DNA profiles corresponding to a criminal affair are recorded in the NDD</td>
</tr>
</tbody>
</table>

Table II. Types of identification numbers
To substantially modernize the exploitation of forensic DNA identification in Belgium, Framework 1 was recently amended (Moniteur belge, 2011, 2013, 2014). We refer to that last set of legislative amendments as “Framework 2”. In essence, Framework 2 specifies aspects that were left undefined in Framework 1, reinforces the judiciary control and data protection, facilitates the exploitation of forensic DNA identification, imposes a greater scientific and administrative standardization, stimulates the international exchange of DNA data, establishes a supervision of the person IDs, enhances the type of intelligence that can be gathered from the NDD and modernises the deletion of DNA profiles from the NDD (Renard et al., 2013).

**Records management issues**

*Expected qualities*

Before we go on and describe the difficulties that have been noticed with managing DNA-based judicial records, we should reflect on what should be expected from them. As it is mentioned in standards of the International Standard Organisation (e.g. ISO 15489; ISO 16175), and discussed in the literature (Committee on Current Records in an Electronic Environment, 2005; Joseph et al., 2012), records have to satisfy certain properties if they are to be used as evidence.

The Committee on Current Records in an Electronic Environment (2005) refers to two characteristics from the Guide for Managing Electronic Records from an Archival Perspective (International Council on Archives, 1997): authenticity (i.e. the persistence over time of the information the records is supposed to hold) and reliability (i.e. the trustworthiness of records as being able to serve as evidence). In their workbook “Electronic Records: A Workbook for Archivists”, the Committee also refers to two other

<table>
<thead>
<tr>
<th>Step</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opening or closing a criminal affair</td>
<td>A case is opened: alerted by the police forces, or from his or her own initiative, a magistrate belonging to one of the judicial districts takes charge of a criminal affair. Alternatively, the case is closed: either because the case is dropped or dismissed, or a person has been convicted and his or her reference profile must be established and recorded in the NDD.</td>
</tr>
<tr>
<td>Sampling biological material</td>
<td>Biological material is sampled by a police force or a judiciary expert (e.g., a medical examiner) on the crime scene, from a suspect, a convicted offender or a victim. It is then stored in the court registries and sent to a DNA laboratory for expertise.</td>
</tr>
<tr>
<td>DNA analysis</td>
<td>An authorized DNA laboratory analyses the sample and succeeds in establishing DNA results, which are reported to the magistrate as well as to the NDD department.</td>
</tr>
<tr>
<td>NDD management</td>
<td>The NDD department opens a case to relate criminal affairs and the DNA profiles with each other, compares the profiles to the NDD, reports the results of the comparison to the magistrates and stores or delete some profiles</td>
</tr>
</tbody>
</table>

Table III. A schematic view of the DNA-based judicial records management process

---

RMJ 25,3

---

274
closely related records characteristics that were subsequently defined in ISO 15489-1 Part 1 (*International Standards for Organisation, 2001*): integrity (i.e. the characteristic of a record that is complete and unaltered) and usability (i.e. the ability to find and properly exploit a record).

Such qualities for DNA-based judicial records are crucial as a lack thereof may affect the pursuit of the judicial truth by them being either unusable (e.g. by not helping in making any progress) or even actively damaging (e.g. by leading an investigation in a spurious direction). It is therefore essential that all necessary safeguards are put into place to reduce the risks associated with such possible defects. In the following discussion, we use our own concepts of completeness, correctness, traceability and usability – which express the ideas discussed by the Committee and defined in ISO 15489-1 in a slightly different way.

**Completeness.** Records must be complete so that they contain the information that is formally expected from them. For instance, a reference profile must be associated to a person ID referring to a known person, a profile must be associated to a criminal affair ID, all the analysed loci should be part of the DNA profiles sent to the NDD, etc.

**Correctness.** They should be correct in the sense of being factually correct (i.e. no conflicting with the truth). A record could be complete but contain incorrect information. For instance, a DNA profile that is recorded with an incorrect sample ID or that presents an error in an allele (e.g. a “1” instead of a “2”) is incorrect. Reliable records need to be both complete and correct.

**Traceability.** Records must be traceable in the sense where we have a proof of their origin, a comprehension of the process that created them, knowledge of the true identity of the person who has contributed to their content, etc. This is the records management concept of authenticity. Along with the notions of completeness and correctness, the notion of traceability is connected to the notion of trust. For example, a record may be complete and correct, but the absence of proof that only authorized staff has had access to it may cast doubt on its trustworthiness. As another example, we should have proof that a record that is supposed to have been created by person X of department Y, has actually been created by person X of department Y.

**Usability.** A record must ultimately be usable for its intended purposes. A record could be complete, correct and traceable but still be unusable. Records are expected to be usable so that they fulfil the objectives, and protect the interests of, the relevant stakeholders. In the case of judicial records, they would be expected to genuinely support the exposure of the truth, the pursuit of justice, the maintenance of a state of law, etc. For example a record could become unusable if, because of lack of organization of the institution using it, it cannot be retrieved or communicated in due time. As another example, a record could be deemed illegal because it is based on information that should not have been recorded in the database or that should have been deleted.

**A divided world**

As we suggested in the subsection “Framework 1” when we defined DNA-based judicial records (see also Table III), many difficulties with such records can be viewed as issues that arise due to the divisions that naturally exist between the various stakeholders of the criminal security and justice system within Belgium (i.e. the judiciary, the police forces, the DNA laboratories and the NDD department) and also between these national stakeholders and their counterparts in partner countries.
One way of looking at these difficulties is that they prevent the proper linking of the various pieces of information that are created, transmitted and used by these stakeholders and that they form together the DNA-based records. Figure 1 is a simplified scheme that illustrates the idea that links must be drawn between different pieces of information, hereby implying that any link or linked information that is absent or incorrect may induce issues of completeness, correctness, traceability or usability of the records. Examples of these issues are provided in the “Issues and practices” section. For the sake of clarity, important stakeholders such as the DNA laboratories and the police forces are not visible on Figure 1.

In the example of Figure 1, Magistrate α is responsible for Criminal affairs 1 (Crim 1) and 2 (Crim 2), whereas Magistrate β handles Criminal affair 3 (Crim 3). Criminal affair 1 corresponds to administrative case NDD 1 opened in the NDD department. Criminal affair 2 corresponds to NDD 2 and Criminal affair 3 corresponds to two NDD administrative cases: NDD 3 and NDD 4. NDD 1 contains three profiles P1, P2 and P3, whereas NDD 2 contains P4 and P5, NDD 3 contains P6 and NDD 4 contains P7. Profile P3 is a stain profile and it matches with convicted offender reference profile P4. Therefore, they both refer to a known person, a convicted offender who is assigned Person ID 1. Profiles P5 and P6 are both stain profiles. They match but refer to a yet unidentified person (who therefore does not have a Person ID) and form an unidentified cluster of profiles.

Issues and practices
In this section, we review in some details the most important issues that arose during the course of developing practices around DNA-based judicial records management.
Overall, these practices have to be understood as attempts to satisfy legal constraints or
to fill the void left by legal texts in particular areas to solve practical records
management objectives of enforcing the completeness, correctness, traceability and
usability of records. Whereas most of them satisfy European standards (ENFSI DNA
Working Group, 2014), they are also specific to the Belgian context.

Non-automaticity and omission. In Framework 1, the acts of performing a DNA
expertise, transmitting the result to the NDD, and recording or comparing the resulting
profile to the NDD were separate acts that needed to be authorized and triggered by
corresponding written judiciary requests. In other words, they were not automatically
performed. As a consequence, due to human mistakes, lack of organization or lack of
apparent usefulness perceived by the magistrate, these acts were sometimes omitted.
For instance, some biological materials that were sampled on a crime scene may never
be analysed, some analysed samples on which a good DNA profile was, however,
obtained may never reach the NDD, and some results that reached the NDD department
may never be recorded or compared. Although this lack of automaticity ensured the
spirit of Framework 1 of a tight control of forensic DNA by the judiciary, the induced
lack of completeness in some records may have prevented important evidence to be
exploited by the judiciary (Renard, 2008; Renard et al., 2013) – a state of affairs that is
addressed and largely solved by Framework 2. Moreover, a greater automaticity of the
vital operations concerning the creation and use of the records not only induces more
complete records but also participates to the reduction of delays in pursuing the judicial
truth. In other words, the records become potentially more rapidly available, and
therefore more “usable”.

Error-prone and inefficient manual operations. A related problem to the
non-automaticity of certain acts is the existence of error-prone and inefficient manual
operations. Under Framework 1, the NDD department would open a case for each
judiciary request, or when DNA profiles were transmitted by a laboratory before the
written judiciary request had been received. Administrative work was therefore
necessary to reconcile request(s) and DNA profile(s) within a NDD case, which increased
the number of communications between the stakeholders and also opportunities for
error prone manual operations. For instance, the NDD department had to contact the
magistrate when a judiciary request was missing while DNA data had been received or,
conversely, had to contact the DNA laboratory when data were missing, whereas a
judiciary request asking to process the data was available.

Moreover encoding mistakes induces additional administrative work to recover from
them. For example, a judiciary request would say that a comparison of a DNA profile
under criminal affair X will have to be performed and that later this profile is sent by the
laboratory under the wrong reference “criminal affair Y” (instead of “X”). These two
pieces of information will lead to open two distinct files (NDD case X and NDD case Y)
that will unduly wait for their respective judiciary request or data for reconciliation,
until somebody notices the error (e.g. after communicating with the stakeholders) and
concludes that they belong to one NDD case instead of two. This state of affairs
considerably slowed down the forensic process and consumed precious resources that
may have been used differently.

With Framework 2, the process of transmitting the DNA profile to the NDD and their
comparison does not need an explicit judicial request anymore, hereby solving the issue
of non-automaticity. It also increases the number of profiles that are inserted in, and
compared to, the NDD. Although it does not mean that all profiles are sent to the NDD, as a magistrate may still decide not to allow such a transfer and that the DNA laboratories will select profiles to be transferred to the NDD department based on the inclusion criteria of the NDD. Although this simplification induces a greater automaticity in the NDD feeding process, as the DNA data received from the laboratories are the only source of information received by the NDD, there is no alternative source of information available to the NDD to check the quality of the data sent by the DNA laboratories. This shows that this desire to improve completeness and usability of the records may create additional problems of correctness that need to be addressed.

Electronic issues. Through Framework 2, the DNA data must be transmitted through secured electronic channels – to the exclusion of post mail transmissions – making the process of inserting the data in the NDD faster (i.e. the record becomes more usable) and less subject to error (i.e. the record is likely to be more correct). Taking advantage of the electronic communication enforced by Framework 2, considerable efforts have been put in developing automated process to import the data transmitted by the laboratories into the administrative system of the NDD. These two steps of electronic transfer from a laboratory to the NDD department (the “transmission step”) and the import of the received data into the administrative system (the “import step”) are, however, associated with questions on the correctness of the transfer and the traceability of the records. For instance, are the channels used to transfer the information sufficiently secured or can they be easily intercepted (and perhaps also altered) by unauthorized parties? Do the import processes faithfully record the data in the administrative system? Ultimately, these questions concern the four characteristics of completeness (e.g. some information could be lost during transmission), correctness (e.g. an altered record could become incorrect), traceability (e.g. if data have been created by laboratory X, how are we sure that they have only been accessed by laboratory X and not by an unauthorized third-party?) and usability (i.e. ultimately, an incomplete, incorrect or untraceable record cannot be used as proper evidence).

The transmission step is secured by an electronic signature and protected by the encryption and decryption program called “Pretty Good Privacy” (PGP). This program works with public and private keys (Garfinkel, 1995). The DNA laboratories and the NDD department have their own public keys that they have exchanged between each other through a non-electronic fashion. The physical character of that exchange ensures the validity of the public keys. Moreover, each user has a private key which is confidential and only used by the owner through a secret password. This system guarantees the security of the exchange between the DNA laboratories and the NDD department. The security involving the users/operators within the NDD department obeys different rules, and is the object of ongoing development.

The aforementioned import step is an important part of the automaticity of the process that leads to the creation of DNA judicial records. A log is automatically updated to describe the operations performed during the import process. The log is consulted to detect errors (e.g. an error can occur when the data sent to the NDD department does not respect the official template) and is used to validate the process. The genetic profiles sent by the DNA laboratories are part of Excel files sent as encrypted attachment of emails. During the import step, the structure of the Excel files (composed of three spreadsheets), the name of the columns (i.e. the variable names) and
the expected data types are compared to a specification template and the internal consistency between the three spreadsheets is checked. The import step has been validated by manual inspection of the input (i.e. the data to be imported) and the output (i.e. the data that has been imported), in situations where the input was ideal and in situations where the input was faulty (e.g. incorrect data types, missing data, incorrect variable names, inconsistency between spreadsheets).

Result quality. The core business of the NDD department is to report comparison results to the judiciary. The quality of the results as well as the quality of the DNA profiles and other administrative data that led to obtaining these results are of key importance. Because Framework 1 did not specify the scientific nature of the acts to be performed to produce correct and usable NDD results, some decisions had to be taken by the NDD department itself, in cooperation with the DNA laboratories. To balance the ability of the NDD to discover true links between DNA profiles and filter out unreliable results, pragmatic decision rules were progressively designed to determine what profiles could be stored in the NDD (“inclusion rules”), and what matches could be considered (“matching rules”) before possibly reporting them. Some of these rules were encoded in Framework 2.

Deletion of DNA profiles. Because reference profiles and stain profiles may be recorded in the NDD, and that their presence (or absence) in this database has strong implications for future results and investigations, specific procedures have been designed to regulate their disposal. For instance, stain profiles may be erased when they are deemed no longer useful. Most of these decisions are taken on practical and scientific grounds. For instance, when the NDD department has identified a stain by matching it to a reference profile registered within the same judicial case, it asks permission from the judiciary to delete the identified stain. Moreover, profiles which are considered for deletion are those that are recognized as being duplicates of already existing records belonging to the same judicial case, or that do not satisfy the inclusion rules anymore because these rules have been adapted. The adaptation of these rules is necessary to reduce the number of unreliable matches (i.e. false positives) that result from increasing the number of recorded DNA profiles and that is aggravated by a greater proportion of DNA profiles of poor quality.

Balancing false-positives and false-negatives. Matching rules involving the count of the number of matching loci, probability calculation or complementary DNA laboratory checks determine when a match between two profiles is considered as “close to certainty”, “highly probable with some reason for doubt”, “possible match” and “match disconfirmed after further inspection”. These rules attempt to balance the rate of false-positives (the proportion of matches that are found but are not true) and the rate of false-negatives (the proportion of matches that are not found but are true). We do not currently have an estimate of the proportion of false positives. Given the matching rules that are used, this proportion is thought to be extremely small if not null. As far as the false-negatives are concerned, some of them could result from encoding errors in the DNA profiles. To detect these errors, a special matching strategy that authorizes one mismatch between loci is applied several times a year to retrieve matching profiles that would be unnoticed otherwise.

Misleading matches. Besides a match being false, it may be correct but still be misleading. For example, profiles may result from contaminating the crime scene with biological material from the sampling kit, the law-enforcement agent sampling the scene
or the DNA expert analysing the sample, which can lead to matches that are not related to criminal acts per se and that are therefore misleading (the profile being found repeatedly on different crime scenes). We do not have a reliable estimate of the rate of contamination because feedback on the existence of contaminations is usually not directed toward the NDD department. When the NDD learns that a DNA profile is the result of a contamination, however, a list of cases is updated and the concerned magistrates are requested to decide how to address the situation (e.g. by deleting the spurious profiles). Another example is about the correct matches that can be found between profiles of persons who do not necessarily play a role in the offence being investigated. The DNA-based evidences that are reported by the NDD always need to be carefully interpreted by the judicial authorities by taking into account not only the NDD report but also the report of the DNA laboratory and the context of the criminal affair.

Useful reporting. Observing a link between DNA profiles (a positive result) or observing no link (a negative result) is usually reported to the judiciary. In Framework 1, all judicial requests had to be answered formally by post mail, even in case of negative result or in the absence of any result (i.e. no profiles could be obtained), which could be perceived as an undue task. Given the lack of resources of the NDD department at the beginning of its installation, it was decided to generally focus on performing the comparisons and reporting the positive results – temporarily leaving aside formal replies involving negative results or absence of results. With Framework 2, the procedure has been simplified in the spirit of limiting the administrative burden that resulted from reporting all types of results.

When a positive link is established between two profiles (or more), the profiles are said to belong to the same cluster. A cluster is, therefore, a set of matching profiles likely to originate from the same person and that corresponds to several criminal affairs (for an example, see Figure 1). Each additional positive result will increase the size of the cluster, leading to a new official version of the corresponding cluster document that is mailed to all the magistrates involved in the cases – which, as the cluster grows over time, can become a laborious administrative task. Under Framework 2, the procedure gained in efficiency. The first time a profile from a case X matches with the profile of another case Y (thereby forming a cluster), the magistrates of cases X and Y are personally informed. In a future match with a new profile from case Z, only the magistrate from case Z will be personally informed. A general mail will however be sent to a supervisory judicial authority who will dispatch the information to the magistrates if necessary.

Absence of feedback. Besides these considerations of unnecessary NDD workload, genuine questions on the usefulness of the intelligence provided by the NDD can be raised. From a point of view of evaluating public policies, there is currently no mechanism to evaluate the relevance of the delivered intelligence (e.g. is it crucial to solve affairs?) and the NDD receives no feedback on the usefulness of its results beside the occasional reading of journalistic accounts or participation as expert witness in criminal trials. From a cost-benefit point of view, this begs the question of whether the resources spent for forensic DNA identification are reasonable deployed. Let’s review two examples of situations where the resources could be used differently.

Potentially non informative positive results. When a person has already been identified through matching stain profiles with a reference profile – and therefore forming an identified cluster of profiles – and that new matching profiles are
subsequently added to the cluster, formal reports still need to be sent to all (potentially numerous) magistrates concerned by that individual, although an identity is available. Similarly, when the magistrate decides to group certain cases together (e.g. because of the similarity of the modus operandi observed in these various cases), it would seem more efficient to send a single mail to that magistrate when modifications concerning these cases occur, rather than sending multiple mails for each single modification. The legal constraints set by Framework 2 and the absence of feedback from the magistrate who decided to group these cases, prevent the NDD to proceed more efficiently in terms of communication policy. As another example, by lack of feedback, the NDD is not aware of when a criminal affair has been closed (e.g. because an identity is known to the judiciary but not available to the NDD), and remains anyway legally obliged to report positive results, regardless of the status of the case. It seems that more efficient ways to communicate on these kinds of results could be imagined (e.g. by developing a feedback system from the judiciary to the NDD, and a prioritization of the reporting).

Towards greater intelligence and cooperation. A particular type of intelligence which caught the attention of the NDD manager at an early stage of the development of this institution is the ability to draw links between clusters (i.e. offenders who have left their DNA profiles on several crime scenes) by relying on the criminal affairs in which the profiles of different offenders occur together (e.g. because they co-offended). For instance, in Figure 1, Person 1 left profile P3 in Crime 1 and profile P4 in Crime 2 and Person 2 left P5 in Crime 2 and P6 in Crime 3. Therefore Person 1 and Person 2 co-offended in Crime 2 – hereby forming a network of clusters. Under Framework 1, magistrates are only informed of updates of the clusters that directly concern them (because they contain a profile associated to a case they supervise). Therefore, they are not informed when a “co-offender cluster” has been changed. In the example, this means that Magistrate α is informed when DNA profiles are “added” to both Person 1 and Person 2 but that Magistrate β is only informed of modifications to Person 2. Such information could however be potentially helpful to all magistrates – especially given the fact that it would not be available otherwise, due to the separation existing between the criminal cases as well as between the judicial districts that handle them. The opportunity to obtain greater intelligence (i.e. records usability) and facilitate the cooperation between the magistrates is opened by Framework 2 and is under investigation (Jeuniaux et al., In Press).

Incoherence and duplication. The profiles recorded in the NDD are currently stored and managed by the CODIS (“Combined DNA Index System”) software, while the administrative activity generated by the NDD business (e.g. receiving a profile, creating a file) is logged in a distinct electronic administrative system (hereafter “ADMIN”). The CODIS software is a product developed by the American FBI that is used in the majority of European Union member states as it is in the USA whereas ADMIN is an in-house developed database management system.

Database coherence. The fact that the NDD is composed of two separate systems creates an opportunity for database incoherence. For instance, a profile might be stored in CODIS but not recorded in ADMIN and vice versa. Also, since the only information that can relate a profile in CODIS and its representation in ADMIN is the sample ID, a mere spelling mistake in the sample ID will have an impact on the management process. To tackle this problem, a higher integration of ADMIN with CODIS has been developed.
To ensure the coherence between the two systems, data management routines are carried out on a regular basis.

Stakeholder coherence. Other types of incoherence, such as issues of duplication, might reveal deeper problems that may pertain to the NDD department or other spheres of the justice system. We define *duplication* as the suspicious repetition of the “same” information (which is purposefully a broad definition that we only aim to explain by way of examples). In an ideal world, there would be a one-to-one (i.e. perfectly coherent) relationship between a person, a person ID and a reference profile – as shown in Panel (1) of *Figure 2* – but this is not what has been observed in practice.

Under Framework 1, the 33 judicial authorities had their own Person ID assignment system and nomenclature, and their Person IDs were usually transmitted to the NDD along with the results sent by the DNA laboratories. It was therefore possible to sample several times a same person who would be convicted in different judicial districts, and assign different person IDs without noticing it. Actually, for some time, a few authorities misinterpreted the purpose of the system as one that required the establishment of a new DNA profile for each new conviction of a same individual, as if the profiles of the NDD constituted parallel criminal records. The centrality of the NDD and its ability to find matching profiles turned out to be a key to detect such incoherencies. In Framework 2, this problem was explicitly tackled by creating a central system of delivery and supervision of Person IDs – which is still being tested as it attempts to detect potential remaining incoherencies.

The other panels of *Figure 2* illustrate several coherence problems. Panel (2) shows that a person might have one person ID and have his or her reference profile made more than one time. Panel (3) shows that a person could use aliases (e.g. because pretending to be different persons would help preventing the accumulation of convictions on a single head, hereby evading more severe convictions) unbeknownst to the law-enforcement and judiciary (e.g. by lack of organization making it difficult to know when a person had been already convicted) and be therefore attributed more than one person ID. Under Framework 1, this possibility of using aliases was facilitated by the logic under which the establishment of DNA profiles was managed. In Framework 2, the control of the identity of the individuals by other biometric means addressed this issue. Panel (4) illustrates that two identical twins may receive different person IDs (as they should, as they are different persons) and distinct recorded profiles (because they have been sampled at distinct occasions) that turn out to be a perfect match (as would be expected). Panel (5) shows two different persons having different profiles but sharing the same...
person ID, which defeats the purpose of a person ID (which, per definition, is supposed to be unique for each distinct person). This last example is likely one of the worst case scenarios. This situation, as well as those depicted in the other panels, are the subject of much attention and cross-checking procedures by the NDD department and the judiciary organism that supervises the assignment of the person IDs.

Conclusion and discussion

Forensic DNA identification technology is an important instrument that is used to establish facts which it would be hard to dispute in criminal cases, to contribute to the detection of transnational crimes, to solve false identities, etc. But it also creates new problems, which are especially relevant from a records management point of view. The technology seeks to exploit a biological identity but is itself dependent on other types of identities and records holding personal or critical information, whether we have in mind the administrative records corresponding to putative criminals or the staff responsible for the creation, transmission or management of the records.

In this paper, we presented issues of records management through the process of DNA-based judicial records management. For sake of space, important aspects have not been treated in this paper. They mainly concern the challenges posed by the greater needs for automation and standardization of the records management process. As far as automation is concerned, there is more to be said of the process that guarantees the validity of computer-based operations, the need to satisfy legal requirements for using paper traces and the necessity to imagine disaster recovery mechanisms that allow functioning in case of electronic breakdown (Jonckheere, 2013). From the standardization point of view, one may wonder, for instance, about the standardization of other aspects of the DNA-based judicial records such as the standardization of reports. Magistrates often do not have the time or the scientific knowledge to read and understand detailed scientific reports. This situation encourages the production of reports that are as succinct as possible, which begs the challenging question of how we can satisfy both objectives of scientific correctness and operational usability. A similar but probably more problematic question concerns the standardization of the part of the reports that deal with the statistical reasoning on the basis of which crucial decisions are taken (Jonckheere and Renard, 2014).

Finally, this paper has three main practical implications. First, it opens up the debate on a critical process of general interest to a democratic society that may rightly be concerned with what safeguards are put into place to protect civil liberties in the face of using powerful biometric tools. Second, it shows successive problems and legislative or organizational solutions surrounding the management of DNA-based judicial records that should be of interest to DNA database and criminal justice professionals who may be encountering similar issues. Third, it shows how the distributed character of the DNA-based judicial records (i.e. the fact that they depend on various divisions or stakeholders) influences the objectives of making these records complete, correct, traceable and usable. These four characteristics are useful concepts to take into account in reviewing the qualities and issues of records management environments of this nature.

References


Renard, B. (2008), Ce que l'ADN fait faire à la justice: Sociologie des traductions dans l'identification par analyse génétique en justice pénale (Doctoral en criminologie), Université catholique de Louvain, Louvain-la-Neuve, 5 November.


**Author affiliations**

Patrick P.J.M.H. Jeuniaux, National Forensic DNA Database, Nationaal Instituut voor Criminalistiek en Criminologie/Institut National de Criminalistique et de Criminologie, Brussels, Belgium

Bertrand Renard, Criminology Department, Nationaal Instituut voor Criminalistiek en Criminologie/Institut National de Criminalistique et de Criminologie, Brussels, Belgium

Leen Duboccage, National Forensic DNA Database, Nationaal Instituut voor Criminalistiek en Criminologie/Institut National de Criminalistique et de Criminologie, Brussels, Belgium

Séverine Steuve, National Forensic DNA Database, Nationaal Instituut voor Criminalistiek en Criminologie/Institut National de Criminalistique et de Criminologie, Brussels, Belgium

Caroline Stappers, Criminology Department, Nationaal Instituut voor Criminalistiek en Criminologie/Institut National de Criminalistique et de Criminologie, Brussels, Belgium

Inès Gallala, Faculty of Law and Criminology, Vrije Universiteit Brussel, Brussels, Belgium

Sabine De Moor, Institute for International Research on Criminal Policy, Ghent University, Ghent, Belgium

Alexia Jonckheere, Criminology Department, Nationaal Instituut voor Criminalistiek en Criminologie/Institut National de Criminalistique et de Criminologie, Brussels, Belgium

Benjamin Mine, Criminology Department, Nationaal Instituut voor Criminalistiek en Criminologie/Institut National de Criminalistique et de Criminologie, Brussels, Belgium

Beatrice Vanhooydonck, Nationaal Instituut voor Criminalistiek en Criminologie/Institut National de Criminalistique et de Criminologie (NICC/INCC)

Morgane Kempenaers, Nationaal Instituut voor Criminalistiek en Criminologie/Institut National de Criminalistique et de Criminologie (NICC/INCC)

Christine De Greef, Nationaal Instituut voor Criminalistiek en Criminologie/Institut National de Criminalistique et de Criminologie (NICC/INCC)

Pierre Van Renterghem, Europol, The Hague, The Netherlands, and

Vanessa Vanvooren, Nationaal Instituut voor Criminalistiek en Criminologie/Institut National de Criminalistique et de Criminologie (NICC/INCC)

**About the authors**

Patrick P.J.M.H. Jeuniaux has a doctoral degree in psychology and degrees in statistics and artificial intelligence. He conducts research in criminology and forensic science at the National Forensic DNA database department at NICC/INCC. Patrick P.J.M.H. Jeuniaux is the corresponding author and can be contacted at: patrick.jeuniaux@gmail.com

Bertrand Renard has a doctoral degree in criminology and a degree in law. He is a guest lecturer in criminology at Université catholique de Louvain, and a research fellow at the Interdisciplinary Research Center on Deviance and Penalty (CRID&P). He conducts research on the use of forensic expertise in the criminology department at NICC/INCC.

Leen Duboccage has a master’s degree in biochemistry. She is a database expert in the National Forensic DNA database department at NICC/INCC. She often represents Belgium at the CODIS meeting.
Séverine Steuve has a doctoral degree in biomedical sciences. She is a database expert in the National Forensic DNA database department at NICC/INCC. She is in charge with supervising quality in the department.

Caroline Stappers has a master’s degree in criminology. She is a researcher in the criminology department of the NICC/INCC and an affiliated junior researcher in the Leuven Institute of Criminology (LINC) at Katholieke Universiteit Leuven.

Inès Gallala has master’s degrees in criminal law and international law. She is a doctoral researcher at Vrije Universiteit Brussel.

Sabine De Moor has a master’s degree in criminology. She is a doctoral researcher in the Institute for International Research on Criminal Policy (IRCP) at Ghent University.

Alexia Jonckheere has a doctoral degree in criminology and a master’s degree in law. She conducts research on managerial justice in the criminology department at NICC/INCC.

Benjamin Mine has a doctoral degree in criminology. He has conducted research on criminological statistics and judicial databases in the criminology department at NICC/INCC.

Beatrijs Vanhooydonck has a doctoral degree in biology. She is a database expert in the National Forensic DNA database department at NICC/INCC.

Morgane Kempenaers has a doctoral degree in molecular biology and microbiology. She is a database expert in the National Forensic DNA database department at NICC/INCC.

Christine De Greef has a doctoral degree in science (chemistry, option biochemistry). She was a database expert in the National Forensic DNA database department at NICC/INCC. She now works in the Biomedical Research Institute at Universiteit Hasselt.

Pierre Van Renterghem has a doctoral degree in chemistry. He is the former custodian of the National Forensic DNA database at NICC/INCC. He was a senior specialist at Europol, where one of his missions was to coordinate tasks related to forensic science. He is now General Manager of WELBIO, an inter-university life sciences research institute based in Wallonia, Belgium.

Vanessa Vanvooren has a doctoral degree in biomedical sciences. He is the head of the National Forensic DNA department and the current custodian of the database at NICC/INCC. She is the chair of the DNA database group of the European Network of Forensic Science Institutes (ENFSI).