Journal of Investigative Psychology and Offender Profiling
J. Investtig. Psych. Offender Profil. **9**: 13–29 (2012)
Published online in Wiley Online Library (wileyonlinelibrary.com). **DOI**: 10.1002/jip.1353

How Long Do Offenders Escape Arrest? Using DNA Traces to Analyse when Serial Offenders Are Caught

MARRE LAMMERS*, WIM BERNASCO and HENK ELFFERS

Netherlands Institute for the Study of Crime and Law Enforcement (NSCR), Amsterdam, The Netherlands

Abstract

Why is it that some serial offenders are arrested quickly and others only after a long period, or never at all? What characterises offenders who continue to escape arrest despite their continued involvement in crime? To be able to answer these questions, arrested (identified) offenders must be compared with never arrested (unidentified) offenders. In this paper, data from the Dutch DNA database are used to assess which characteristics of the criminal behaviour of unidentified offenders influence the probability that they will subsequently be arrested. DNA data offer a unique opportunity to compare crime series of identified offenders with crime series of yet unidentified offenders. Using the Cox proportional hazards model, we tested whether the number of crimes committed, offence specialization, and offence seriousness affect the probability of arrest of serial offenders. Results showed that as an offender commits more crimes, the probability that he will be arrested increases and that offence specialization decreases the probability of arrest. Another conclusion drawn is that DNA traces offer unique opportunities for criminological research. We discuss the limitations of this new data source and make suggestions for future research using DNA traces and for future research that might improve the current study. Copyright © 2012 John Wiley & Sons, Ltd.

Key words: clearance; serial offenders; DNA traces; unidentified offenders; Cox proportional hazards model; hazard ratio

In many western countries, only about 20% of all reported crimes lead to arrest (Dodd, Nicholas, Povey, & Walker, 2004). This percentage, the clearance rate, is seen as an important measure of police performance (Addington, 2006). But it is more than just that. Solving crimes is crucial to maintain the legitimacy of the criminal justice system and the effectiveness of its sanctions. Low clearance rates have a negative influence on public confidence in the police and in the criminal justice system (Jiao, 2007; Litwin & Xu, 2007; Riedel & Jarvis, 1999; Roberts, 2007). Furthermore, because certainty of punishment is important for the effectiveness of both general and specific deterrence (see, e.g.,

*Correspondence to: Marre Lammers, Netherlands Institute for the Study of Crime and Law Enforcement (NSCR), PO Box 71304, 1008 BH Amsterdam, The Netherlands.
E-mail: MLammers@nscr.nl

Blumstein, Cohen & Nagin, 1978; Nagin, 1998), unsolved crimes make offenders expect they can continue to commit crimes and still escape justice (Paré, Felson, & Ouimet, 2007). In fact, there must be quite a number of offenders who have committed multiple crimes, sometimes even very serious crimes, but who have never been arrested and brought to justice for any of these crimes. From this perspective, it is not only important to understand the factors that influence whether individual crimes are solved but also to understand what characterises offenders who continue to escape arrest despite their continued involvement in crime.

Without exception, prior studies have addressed clearance as an issue pertaining to individual crimes and have therefore taken the *crime* as their unit of analysis and its *clearance* (or the time until clearance) as the dependent variable. To address the other clearance issue, that is, the identification of a person as the perpetrator of one or more crimes, the *criminal* must be the unit of analysis, and his[1] *arrest* (or the time until arrest) must be the dependent variable. It is therefore necessary to compare arrested (identified) offenders with never arrested (unidentified) offenders and study the differences between these two groups. The ideal study to answer this issue would be based on the accounts of an unbiased random sample of offenders (some arrested, some not) who would all have perfect memories and be completely honest. But as such samples are not feasible, the present study takes another approach. It utilises a forensic DNA database to reconstruct the offending patterns of serial offenders of whom DNA was secured at two or more crime scenes and relates characteristics of their prior offending to the probability of subsequent arrest.

Another limitation of previous studies on crime clearance is that the large majority deals only with the clearance of homicide cases (see, e.g., Alderden & Lavery, 2007; Davies, 2007; Jiao, 2007; Keel, Jarvis, & Muirhead, 2009; Litwin, 2004; Litwin & Xu, 2007; Paré, *et al.*, 2007; Regoeczi, Jarvis, & Riedel, 2008; Roberts, 2007). The present study is not restricted to a limited set of serious crimes such as homicide and rape but also includes a wide variety of common 'volume' crimes such as burglary and vehicle crime.

The present paper will use the Dutch DNA database to reconstruct the offending patterns of offenders of whom at least two DNA stains are present in the database and to study whether the number of prior crimes, the seriousness of prior crimes, and the level of specialization in prior crimes affect the probability of subsequent arrest. Data from the DNA database offer a unique opportunity to compare crime series of identified offenders with crime series of yet unidentified offenders and to attempt to find out why the latter have remained unidentified thus far. Although there are some caveats associated with selecting crimes from a DNA database as observational units, the main advantage is that it allows us to establish, with almost perfect certainty, links between the crimes of serial offenders, even if they have never been arrested or identified by law enforcement.

After this introduction, we will first discuss previous crime clearance research and justify our hypotheses. Subsequently, we discuss how DNA traces can be used for serial clearance research and explain what data were retrieved from the Dutch DNA database. Next, we discuss the method used (Cox proportional hazards model) to compare crime series of identified offenders with crime series of unidentified offenders and to establish which characteristics of these series affect the probability that the offender is arrested, and the results of the analysis. This paper ends with a discussion of the results and a discussion on using DNA traces for clearance research.

[1]To prevent awkward stylistic constructions, we use male pronouns only when referring to persons.

DOI: 10.1002/jip

THEORY

Previous clearance research

As mentioned in the introduction, previous clearance research focused mostly on the clearance of single homicides. Alderden and Lavery (2007) gave a short overview of these studies. They concluded that most previous studies include at minimum the influence of characteristics of the victim (age, race, and gender), the type of weapon that was used, and characteristics of the incident. Some studies also include characteristics of the community and characteristics of the police force investigating the homicide. Alderden and Lavery (2007) drew three general conclusions from these studies. First, nearly all studies find that the type of weapon is a significant predictor of homicide clearance: when a firearm is used, the homicide is less probable to be cleared. Second, most studies find that the homicide circumstance (e.g. whether a concomitant felony was committed) is a significant predictor of homicide clearance. However, because the measurement and operationalization of the circumstances is different in the various studies, it is difficult to draw specific overall conclusions on these circumstances. Third, the studies have found conflicting results on the influence of victim characteristics on homicide clearance, which might be the result of the different ways the characteristics were measured in these studies.

Because the term ‘clearance’ is mostly used to describe the arrest of an offender for one single crime (that single crime is then considered cleared) and to prevent any misunderstanding, we will use the term ‘serial clearance’ to indicate that when an offender is arrested, all crimes where this offender’s DNA was secured are cleared.

Influence of crime series’ characteristics on serial clearance

This study will focus on how characteristics of crime series influence serial clearance, an issue that to our knowledge has not been studied before. Previously committed crimes can influence the probability of arrest in two ways. First, offender behaviour might change, which changes the characteristics of the ongoing series and thus the probability of arrest. For example, the offender gains criminal experience, or he may become less vigilant after repeatedly having escaped arrest. Second, police behaviour might change on the basis of knowledge they gain about the crimes in a series. When a police department in the Netherlands sends a DNA crime scene sample to the Netherlands Forensic Institute (NFI) for analysis, the NFI reports if the DNA trace matched other crime scene traces and if it matched the DNA sample of a person in the database. On the basis of this expert evidence, the police are then aware of crimes previously committed by the offender, who is known now or still unidentified. If the offender is known, police will direct its efforts in locating and arresting that specific offender. If he is unknown, police may change its investigation strategy, using data from previous now linked offences as well, which may influence the probability of clearance.

We will investigate how specific characteristics of a crime series (the number of committed crimes, the seriousness of the committed crimes, and specialization) influence the probability of arrest. For each of these aspects, we propose a hypothesis.

The first hypothesis is that the more crimes an offender commits, the more probable it is that he will be arrested. Every committed crime carries in itself a certain probability of

DOI: 10.1002/jip

being cleared, on the basis of different aspects of the crime. This probability may very well decrease over time, when a crime remains unsolved up to that moment, but a (possibly rather small) probability of clearance will always remain. Thus an offender who commits multiple crimes faces the weighted probability of arrest of all crimes in his series. As a consequence, we expect that as the number of committed crimes increases, so will the probability that the offender will be arrested.

The second aspect of crime series we study is crime seriousness. Serious crimes, such as homicide, sex offences, and other violent offences, probably get more attention from the police and the criminal justice system than crimes that cause less harm (Heller & McEwen, 1973). Also, because serious violent crimes often involve contact between the offender and the victim, there may be more evidence that increases the probability of clearance. Our hypothesis is that the more serious the committed crimes are, the more probable it is that the offender will be arrested.

The last characteristic of the crime series studied is offence specialization. For an offender to be specialized, there needs to be some degree of repetition of the same offence type. An offender who is not specialised in any crime type is called versatile. When an offender specialises in a certain type of crime, he will become more experienced in committing that type of crime. Research has shown that experienced burglars have superior perceptual and procedural skills, both compared with non-offenders and with novice offenders who just started burgling (Clare, 2010). Topalli (2005, p. 270) stated that '(...) there appears to be adequate support for the notion that criminals develop and maintain some form of expertise related to offending.' Moreover, experienced burglars have a greater awareness of the risks of being caught (Clare, 2010). So, more experience in committing a certain type of crime leads to more expertise, which may lead to a decreased probability of arrest. We hypothesise that (all other things being equal, including the number of committed crimes) a specialized offender has a smaller probability of being arrested than a versatile offender.

DATA

Using DNA traces for serial clearance research

Until now, clearance research has focused on the clearance of single (homicide) incidents, in part because of the inherent difficulties involved in ascribing a series of unsolved crimes to the same unknown offender. If the offender is not identified, it is not possible to know which other crimes he has committed. DNA traces offer a unique opportunity to study the characteristics of crime series even if the offender is unidentified.

Data used to study (homicide) clearance usually are data from official police records and have as such various limitations. Police data provide few if any possibilities to link multiple crimes to one offender, as long as that offender remains unidentified. Consequently, only crime series of identified offenders can be analysed. In order to support police investigations and legal proceedings, linking multiple crimes to one offender when no DNA traces or any other biological evidence (such as fingerprints) is available has been tried on the basis of behavioural similarities (see, e.g., Hazelwood & Warren, 2003; Tonkin, Grant, & Bond, 2008; Woodhams & Toye, 2007). However, because they are consistent and also distinctive, DNA traces enable the linking of multiple crimes to one offender with a far greater certainty and reliability than behavioural evidence. Thus, DNA traces give us the

DOI: 10.1002/jip

opportunity to study crime series of unidentified offenders and to compare them with series of identified offenders in order to assess whether characteristics of these series have an impact on the probability that unidentified offenders become arrested and, consequently, identified. In addition, they allow us to determine how long it takes for an offender to become arrested and to assess the factors that affect how long an offender is able to escape law enforcement.

Data from the Dutch DNA database

The data used in this study were retrieved from the Dutch DNA database. This database was established in 1997 (Schneider & Martin, 2001) and is managed by the NFI.

Two types of DNA profiles are kept in the database. The first type of profiles is subject profiles, obtained from arrested suspects. If the Dutch criminal law allows custody for the committed crime, DNA can be taken from the suspect to compare it with DNA found at the crime scene. This crime scene DNA is the second type of DNA profiles kept in the database. Thus the database contains DNA profiles that are directly taken from a person (the suspect) and DNA profiles that are obtained from the crime scene (e.g. from a cigarette butt).

For every crime scene trace, the database contains information about the crime that was committed, and if a suspect has been arrested, details about this suspect. The database also stores information about which crime scene traces match with other crime scene traces and whether the crime scene trace(s) match(es) the DNA profile of a person (and if yes, with whom). Of every crime scene trace, the database contains the type of crime that was committed, as classified by the police, in which of the 25 police regions the crime was committed, the date on which the crime was committed, and whether a suspect was arrested for the crime.

Before 2001, the Dutch law only allowed DNA to be taken from suspects that were arrested for committing offences for which a prison sentence of 8 years or more could be imposed. A law amendment in 2001 made it possible to collect DNA from suspects that are arrested for crimes for which a prison sentence of 4 years or more can be imposed. Under this amendment, it is allowed to collect DNA from offenders of various high volume crimes, such as burglary. Under the new law, police started to collect DNA traces from crime scenes of high volume crimes. Under the old legal provisions, such crime scene traces would be useless, as they could not be compared with a DNA profile retrieved from a suspect: the law did not permit DNA to be taken from suspects who were suspected of committing high volume crimes. Until 2002, the number of DNA crime scene traces that had been loaded onto the database was on average 690 per year. The law amendment caused an increase in the annual number of DNA traces that are loaded onto the database, which is now on average 4,650 (NFI, 2010). For the present analysis, we use all DNA crime scene traces that were loaded onto the database since the 1st of January 2002 until the 31st of December 2009.

The dataset contains 8,861 offenders, of whom 4,431 left only one crime scene sample behind. We chose to leave these offenders out of the analysis because we want to study the characteristics of crime series (and not of single incidents). The analysed dataset contains 4,430 offenders, who together committed 14,185 known crimes. More than half of these offenders, 2,294, had not been identified by 31 December 2009.

The database contains information not only on whether a crime series is cleared because the offender was identified but also when that happened. Clearance of a crime

DOI: 10.1002/jip

is often defined as the offender being either arrested or convicted (for examples of different definitions, see Regoeczi *et al.*, 2008; Paré *et al.*, 2007). In the present analysis, we speak of serial clearance whenever the DNA profile of the arrested offender matches the DNA traces found at the different crime scenes. The date of serial clearance is the date on which the DNA profile of the offender is loaded onto the DNA database and matches the crime scene traces. Time to clearance is measured as the number of days between the date that the first DNA trace was loaded onto the database and the date of arrest of the offender.

There are two different ways in which serial offenders can become identified. First, by arrest. The offender is arrested for committing a crime where DNA traces were found at the crime scene. Consequently, the police will take his DNA to compare it with the crime scene DNA. The DNA profile is sent to the NFI and is loaded onto the database. In this case, the profile will match the DNA trace found at the crime scene, but it also matches crime scene traces that were collected at crime scenes of previously committed crimes and that were already in the database. Because these traces matched with each other, they formed a series of crimes of one offender. This series is now considered cleared: the offender is identified. Such an arrest does not necessarily happen directly after the last crime was committed, so there might be some time between the last committed crime in the series and the serial clearance date.

A second way in which a crime series can be cleared is a result of a law that was introduced in 2004. This law made it possible to collect DNA profiles of offenders who were previously sentenced to either a prison sentence or a community service for committing a crime. As a consequence of this law, DNA profiles of convicted offenders of whom DNA was not taken before are now taken. The DNA profiles of these known offenders are then loaded onto the database, and it may again match a series of crime scene traces of previously committed unsolved crimes. These crimes were not solved before, because DNA of the offender was not collected before. This way of retrospectively collecting DNA of convicted offenders has led to an increase in cleared crime series.

It is of course possible that an offender commits another crime after the crime series is considered cleared. The DNA profile of the offender remains in the database, so when this offender commits another crime where he leaves a DNA trace behind, this crime will be cleared the moment that the crime scene trace is loaded onto the database. We excluded these later crimes in our research because they were committed after the crime series is considered cleared (according to our definition).

METHOD

Measures of characteristics of crime series

In this paper, we study three characteristics of crime series and their influence on the probability that the offender is arrested: the number of committed crimes, offence specialization, and seriousness. It is important to emphasise that each one of these characteristics can change as the series unfolds over time and is effectively measured at each point in time. In this section, we discuss how we measured the different characteristics.

The first characteristic of a series is the number of crimes that have been committed up to a particular point in time. Because a crime series starts to exist with a crime where a new

DOI: 10.1002/jip

DNA trace is collected (i.e. one that does not match any stain or profile in the DNA database), this variable initially takes on the value 1. When after some time a second DNA trace is added that matches the first, the variable takes on the value 2 and so on. The second measure is offence specialization. Specialization measures the degree of repetition in a series with respect to offence type. We measure specialization taking into account the entire series of crimes that an offender committed up to time *t*. Specialization is also a dynamic variable, as it measures the amount of repetition in the crimes committed up to the moment in time where specialization is measured. The two most widely used measures of specialization are the percentage rule, where specialization is defined by at least 50% of the crimes in the series being one offence type, and the diversity index (Miethe, Olson, & Mitchell, 2006). We use both measures.

The diversity index is a measure of heterogeneity, and applied to criminal behaviour, it reflects the probability that two crimes that are randomly selected from a series of crimes of an offender turn out to be two different types of crime (Mazerolle, Brame, Paternoster, Piquero, & Dean, 2000; Sullivan, McGloin, Pratt, & Piquero, 2006). The diversity index can range between 0, which means absolute specialization (all crimes are of the same type), and 1, which indicates total versatility (no two crimes are of the same type). The maximum attainable score on the diversity index depends on the number of categories that are distinguished (Mazerolle, *et al.*, 2000) and on the number of crimes committed. A feature of the diversity index is that if an offender has a low score (close to 0) and therefore can be called a specialist, it does not indicate the type of crime in which he is specialized. For example, two offenders can both score 0 on the diversity index, with one committing only violent offences and the other only burglaries. Because the risk of arrest may not only depend on crime type consistency but also on crime type itself, we also use the percentage rule: an offender is specialized in a type of crime if more than half of the crimes he committed consist of that particular type of crime. When the percentage rule is used, offences are often coded into certain categories. The most used categories are violent offences, property offences, and public order offences (Mazerolle, *et al.*, 2000; Miethe, *et al.*, 2006), with sex offences either included in the violent category or analysed separately. In this study, the same categories are used with sex offences as a separate category. Public disorder crimes were left out of the analysis, as there were only two offenders in the dataset who committed such a crime. Further, because simple theft, burglary, and theft from car form the large majority (about 85%) of crimes in the sample, we analysed them as separate crimes instead of as part of a single overall ‘property crime’ category. This leaves in total five crime types to be analysed: violent crimes, sex offences, burglary, theft, and theft of or from a car. The category violent crimes consists of assault, extortion, threatening, manslaughter, and homicide.

The third and last characteristic of the crime series is the seriousness of the crimes that are part of the series. The seriousness of crimes is often measured by developing seriousness scales based on surveys of public perceptions, on the opinion of criminal justice professionals (Liu, Francis, & Soothill, 2010; Ramchand, MacDonald, Haviland, & Morral, 2009) or on the maximum sentence that can be imposed by law (Kyvsgaard, 2003; Liu, *et al.*, 2010). In this paper, the seriousness of a single offence is measured as the number of years in prison that can be imposed by Dutch law for this particular type of offence. The maximum numbers of years in prison that can be imposed for committing the different types of crime are 9 years for extortion, 6 years for threatening somebody, life imprisonment or 30 years for manslaughter and homicide (which leads to a maximum prison sentence of 30 years for the violent category), 15 years for sex offences, 9 years

DOI: 10.1002/jip

for burglary, 9 years for theft, and 9 years for theft of or from a car. Seriousness is also a dynamic variable and refers to the crimes committed before the moment in time that seriousness is measured. To calculate crime series seriousness at time *t*, we averaged the crime seriousness of all crimes committed up to time *t*. See Table 1 for an overview of the crime types included in this study and the maximum sentence that can be imposed.

As indicated earlier, our measures typically change over time when the offender commits additional crimes after the initial crime. This implies that the independent variables do not have the same value at every time point (as, e.g., a person's sex would) but that they have values that change over time (an example of this is a person's age). At every point in time that the offender commits a crime, the values of the independent variables (can) change.

The way we measured these time-varying independent variables is as follows. The number of committed crimes is a cumulative measure so that every time an offender commits another crime, the number of crimes increases with one. The seriousness score is a continuous variable and is measured as the mean number of years in prison that can be imposed for the crimes that the offender committed (so far). Because the four specialization variables are dichotomous, their value can only change from 0 to 1, when more than 50% of the crimes committed at that point in time are of the same type or the other way around when the offender becomes more versatile in the types of crime he commits. The diversity index is a continuous variable that ranges between 0 and 1, and can also change at the time a new offence is committed.

Table 2 shows an example of how the variables can change over time. The first offender committed four crimes: two burglaries, a violent crime, and theft. The cumulative number of crimes increases with 1 every time the offender committed another crime. The maximum prison sentence that can be imposed for a burglary is 9 years, so for the first two crimes, the mean seriousness score stays 9. The maximum prison sentence in the violent crime category is 30 years, so the mean seriousness score changes from 9 to 16 ($[9+9+30]/3$). The maximum number of years in prison that can be imposed for theft is also 9 years, which changes the mean seriousness score to 14.25 ($[9+9+30+9]/4$). The diversity index score stays 0 for the first two crimes because they are both burglaries and the offender is therefore totally specialized. It increases as the diversity of the committed crimes increases. As for the specialization variables, the table shows that the first offender is never specialized in violent crimes, and therefore, he scores 0 constantly. For the first three crimes, the score on the burglary specialization variable is 1 because more than 50% of the crimes this offender committed were burglary (100% at the first crime, 100% at the second crime, and 66.6% at the third crime). When he commits the fourth crime, exactly half of his crimes are burglaries, so this offender is no longer specialized in burglary because, as

Table 1. Crime types included in the study

Crime type	Includes	Maximum sentence in years
Violent offences	Extortion, threatening, manslaughter, and homicide	30
Sex offences	Rape, sexual assault, and sexual abuse	15
Burglary	Residential and commercial	9
Theft		9
Theft of or from a car		9

DOI: 10.1002/jip

Table 2. Example of time-varying variables

ID	Offence	Date of offence	# Crimes	Mean seriousness	Diversity index	>50% burglary	>50% violent	Offender arrested	Series cleared
1	Burglary	11.03.2002	1	9	0	1	0	No	
1	Burglary	12.07.2002	2	9	0	1	0	No	
1	Violent	03.21.2004	3	16	0.44	1	0	No	
1	Theft	09.26.2005	4	14.25	0.63	0	0	No	No
2	Violent	01.13.2005	1	30	0	0	1	No	
2	Theft	06.25.2005	2	19.5	0.5	0	0	No	
2	Burglary	06.30.2005	3	16	0.66	0	0	Yes	Yes

stated before, for an offender to be specialized in a certain type of crime, more than 50% of the committed crimes should be the same type of crime. This first offender is not arrested at any time during the study period, so the series is not cleared. If we would have taken only the values of the variables at the end of the study period into account, we would have missed that this offender began as a specialized offender.

The second offender committed three different types of crime. As he commits more crimes of a different crime type, his score on the diversity index increases from 0 at the first offence to 0.66 at his third offence. The first offence is a violent offence, which means a seriousness score of 30. After that, he commits a theft and a burglary, so the mean seriousness score decreases (after the second crime: $(30+9)/2=19.5$; after the third crime: $(30+9+9)/3=16$). This second offender is arrested after he committed the burglary, and his DNA is taken by the police to compare it with DNA traces that were found on the burglary crime scene. The police send the DNA to the NFI where it matches the DNA trace found on the burglary crime scene and DNA crime scene traces that were found at the other two other crime scenes (the violent crime scene and the crime scene of the theft). So, on the basis of the arrest of the offender for the burglary, the entire series of crimes committed by this offender is now cleared.

Descriptive statistics

This paper compares crime series of identified offenders with crime series of unidentified offenders to study whether these characteristics have an influence on the probability of arrest. We will first present some descriptive statistics. However, simple descriptive statistics have two limitations. First, these statistics do not take into account that there is a possibility that offenders who are currently unidentified might be identified in the (near) future. Table 3 shows that one offender in the dataset was only followed for 0.10 months (around 3 days). This offender committed his first offence 3 days before we ended our study period. We were not able to follow this offender for a longer period, and thus, he counts as unidentified. However, it may be that this offender was arrested on (for example) 02 January 2010. In other words, descriptive statistics do not take the possibility of censored durations into account.

A second limitation is that descriptive statistics measure all the characteristics of the crime series at the end of the study period, considering all the crimes of the series together, although these characteristics usually change over time when the offender develops his criminal behaviour.

DOI: 10.1002/jip

Cox proportional hazards model

The two limitations of descriptive statistics and the fact that we do not only want to look at differences between the two groups of offenders but also at which of these differences influence the probability of arrest, make it necessary to perform a survival analysis. Survival analysis studies how long it takes for an event of interest to take place, given that the individual is still at risk for experiencing the event. Survival analysis efficiently utilises censored durations. Applied to this study on arrest of offenders, survival analysis considers the time that it takes for a crime series to be cleared and not just whether a series is cleared or not (Roberts, 2008), and it recognises the possibility that the arrest of the offender can still happen after the end of the study period. It can also take the changing character of the crime series into account.

The outcome measure of a survival analysis is a hazard ratio. The hazard ratio is an indicator of the effect of the independent variable on the hazard (or risk) of the event of interest. The hazard ratio can be interpreted as the change in the hazard (or risk) of experiencing the event of interest that is the result of a one-unit change in the independent, explanatory variable. For instance, the independent variable is 'age in years', the dependent variable is committing an offence, and the model shows that the hazard ratio is 1.20. This means that a 1-year increase in age increases the hazard of committing an offence by 20% (Cleves, Gould, Gutierrez, & Marchenko, 2008). In this paper, the event of interest is the arrest of the offender, and the independent variables are the characteristics of the crime series. The hazard ratio can thus be interpreted as the change in probability of arrest as a result of a one-unit change in the crime series' characteristic of interest.

For this paper, we use a semi-parametric form of survival analysis: the Cox proportional hazards model (Cox, 1972), which has the possibility to use time-varying independent variables, which is essential because we want to take the changing character of crime series over time into account. Another advantage of Cox proportional hazards model, instead of a parametric form of survival analysis, is that it does not require any specification of how the hazard ratio depends on the passage of time (Roberts, 2007).

RESULTS

Descriptive statistics

Table 4 shows the descriptive statistics of the independent variables: the characteristics of the crime series. The first column shows statistics of the unidentified offenders, and the second column shows statistics of the identified offenders. This table shows that identified

Table 3. Descriptive statistics: length of the follow-up period

	Unidentified offenders ($N = 2{,}294$)				Identified offenders ($N = 2{,}136$)			
	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.
Length of the follow-up period (in months)	54.57	*25.74*	0.10	95.38	26.32	*20.90*	0.07	90.85

SD, standard deviation.

DOI: 10.1002/jip

offenders committed on average more crimes (3.5) than unidentified offenders (2.9). The average number of years of imprisonment that can be imposed for the committed crimes is for both groups just over 9 years.

The diversity index is the same for both groups of offenders: 0.21. This indicates that both identified and unidentified offenders tend to be specialized rather than versatile in their offending. Unidentified offenders (1,535 or 67% of the unidentified offenders) are specialized in one of the offence types. More than half of the unidentified offenders are specialized in burglary. Of the identified offenders, 1,644 or 77% are specialized. Again, more than half of the identified offenders are specialized in burglary. When looking at the descriptive statistics, the characteristics of the crime series of identified and unidentified offenders seem to be quite similar.

Table 3 contains information on the length of the period that the offenders in the data were 'followed'. This is the number of months between the first committed crime of an offender and either the date of arrest of that offender or the end of the study period (31 December 2009). This table shows that the unidentified offenders were followed for around 54 months (approximately 4.5 years), whereas the identified offenders were followed for rather more than 2 years (26 months).

Figure 1 shows the (Kaplan–Meier estimate of the) survival curve, a descriptive statistic of the dependent variable: the time until arrest. The survival curve avoids the disadvantages of the descriptive statistics and shows the percentage of offenders that remain unarrested over time. The x-axis shows the number of days, and the y-axis shows the percentage of offenders who 'survive' (are not arrested) past time t. The curve shows that after approximately 8 years (3,000 days), around 65% of the offenders have been arrested. This means that around 35% of the offenders will not be arrested within 8 years and will possibly never be arrested. This is a different percentage of unidentified offenders than the percentage (52) that was shown in the descriptive statistics because the estimated survival curve takes into account that the offender can still be arrested after the end of the study period (as described earlier), whereas the descriptive statistics only shows the percentage of offenders that was arrested at the 31st of December 2009, irrespective of how long they had been observed before.

Table 4. Descriptive statistics of crime series' characteristics (until 31 December 2009)

	Unidentified offenders ($N = 2{,}294$)				Identified offenders ($N = 2{,}136$)			
	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.
Committed crimes	2.90	*1.90*	2	22	3.53	*2.74*	2	35
Seriousness	9.23	*1.84*	3.38	30	9.17	*1.39*	4.50	30
Diversity index	0.21	*0.24*	0	0.80	0.21	*0.24*	0	0.75
	N		%		*N*		%	
>50% violent crimes	33		1.44		16		0.75	
>50% sex offences	7		0.31		8		0.38	
>50% burglary	1223		53.31		1320		61.80	
>50% theft	57		2.48		97		4.54	
>50% theft of/from car	215		9.37		203		9.50	
Not specialized	759		33.09		492		23.03	
Total	2294		100		2136		100	

SD, standard deviation.

DOI: 10.1002/jip

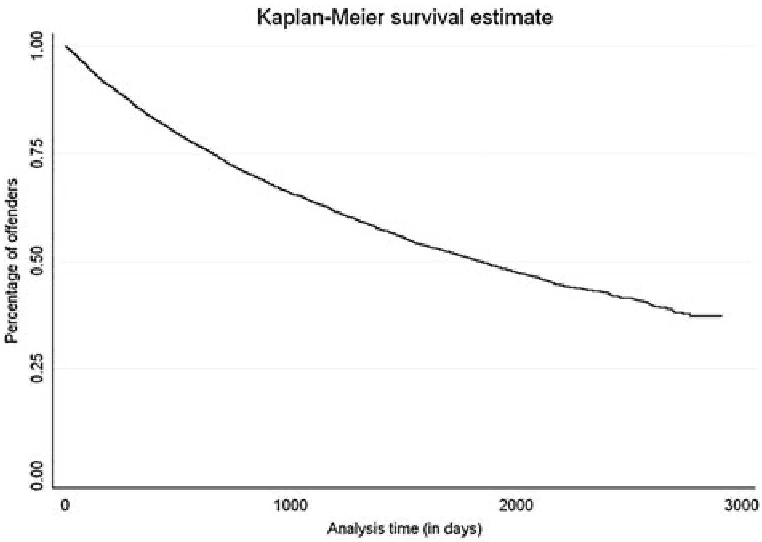

Figure 1. Kaplan–Meier survival estimate.

How do characteristics of crime series influence the probability of arrest?

Two models are analysed using the Cox proportional hazards model. The first model includes the cumulative number of crimes, the mean seriousness score, and the diversity index. The second model includes the crime specific specialization variables and also the cumulative number of crimes and the seriousness score. Because the diversity index measures the same concept (specialization) as the crime specific specialization variables, the diversity index was left out of the second model because it might cause multicollinearity.[2]

Table 5 shows the results of the Cox proportional hazards analysis. In the first model, the cumulative number of crimes has a hazard ratio of 1.19, which means that for every extra crime an offender commits, the probability of him getting arrested increases with a factor 1.19 or 19%. The hazard ratio of the mean seriousness score, which reflects the increased probability of arrest induced by a 1-year increase of average seriousness of previously committed crimes, shows no significant result. The diversity index is a measure of heterogeneity, which means that a high score on this index indicates that an offender is versatile in the type of offences he commits. The hazard ratio of the diversity index (2.84) shows that the more heterogenic the crime series of an offender is, the more probable it is that the offender will be arrested. A one-unit increase in the diversity index means an increase from 0 (absolute specialization) to one (total versatility). Thus, a totally versatile offender is 2.84 times more probable (a 184% increase) to get arrested than a totally specialized offender. So, on the basis of the hazard ratio of the diversity index, the conclusion is that specialized offenders have a smaller probability of getting arrested than versatile offenders.

Because the specialization variables are dichotomous, with a score of 1 meaning that an offender is specialized (more than half of all committed crimes are the same type), we should expect on the basis of the hazard ratio of the diversity index that the hazard ratios of the specialization variables are smaller than 1. A hazard ratio that is smaller than 1

[2]The second model without the diversity index has lower variance inflation factors than the same model with the diversity index included.

DOI: 10.1002/jip

Table 5. Results of Cox proportional hazards model

	Model 1		Model 2	
	Hazard ratio	SE	Hazard ratio	SE
Cumulative # crimes	1.19***	0.01	1.19***	0.01
Mean seriousness score	0.99	0.01	1.01	0.01
Diversity index	2.84***	0.19		
Specialization in				
Violent offences			0.36***	0.09
Sex offences			0.50	0.18
Burglar			0.86**	0.05
Theft			0.43***	0.05
Theft of/from car			0.83*	0.07

SE, standard error. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

indicates that specialized offenders have a smaller probability of arrest compared with versatile offenders, as the results of the hazard ratio of the diversity index show. Indeed, the hazard ratios of four of the crime-type specific specialization variables show this result. When an offender is specialized in violent crimes, the probability that he will be arrested is a factor 0.36 of the probability that a versatile offender will be arrested. The same is true for offenders who are specialized in burglary, theft, or theft of or from a car: these specialized offenders have a smaller probability of getting arrested than offenders who are versatile in their offending (respectively a factor 0.86, 0.43, and 0.83). Specialising in sex offences has no significant influence on the probability of arrest. The results for the cumulative number of crimes and the seriousness score in the second model are almost identical to the results for these two variables in the first model.

CONCLUSION AND DISCUSSION

The results show that 65% of offenders, who left their DNA behind at two or more crime scenes, will be arrested within 8 years. This means that 35% will be arrested more than 8 years after their first DNA-recorded crime, or never. The fact that a rather large percentage of offenders will probably never be arrested is not surprising, considering the low clearance rate described in the introduction of this paper. But why is it that some offenders are arrested swiftly and others only after many years, or never at all?

The results of the Cox model show that the probability of an offender getting arrested increases with the number of crimes he commits. This is in line with our hypothesis, which stated that the more crimes an offender commits, the more probable it is that he will be arrested. So, although a number of offenders will never be arrested, the most prolific offenders who do the most damage to society do have a greater probability of getting arrested. Those who stopped after their second offence are most probable to remain under the radar long or forever.

We also hypothesised that the more serious the crimes of an offender are, the more attention the crimes and the offender will get from the criminal justice system, and the more probable it is that the crime series will be cleared. The analysis shows however that the seriousness of the committed offences has no significant influence on the probability of arrest.

DOI: 10.1002/jip

The hypothesis on offence specialization is that an offender gets more experienced in committing a certain type of crime by committing that same crime multiple times which leads to a smaller probability of arrest. The hazard ratio of the diversity index shows that specialized offenders do indeed have a smaller probability of being arrested compared with offenders that are more versatile in their offending. Looking further into what type of crime an offender is specialized in shows that especially offenders who are specialized in a violent crime or in theft have a smaller probability of arrest, compared with offenders who are versatile in the offences they commit. The result that specialized offenders have a smaller risk of arrest than versatile offenders can be interpreted as evidence that experienced offenders do indeed have more awareness of the risks of getting caught (as described by Clare (2010)) and that they are avoiding these risks.

Concluding, the answer to the main question we raised in the introduction (how do characteristics of previously committed crimes influence the probability of subsequent arrest of serial offenders?) is that characteristics of previously committed crimes do have an influence on the probability that serial offenders are arrested: an offender of whom all committed crimes are the same has a smaller probability of arrest than an offender of whom the committed crimes are different types of crime. Comparing these findings to previous research is difficult because research on clearance focused on single incidents, whereas this paper focuses on serial clearance. Previous clearance research also studied mostly victim or incident characteristics, and this paper shows that besides those characteristics, it is also interesting to study how characteristics of crime series might influence the probability of arrest.

Remarks on using DNA traces for (serial) clearance research

Although our study is innovative as it is probably the first attempt to explain why some criminals are able to escape arrest for long periods, it has limitations. These apply in the first place to the type of data used, DNA traces. As mentioned in the introduction, the ideal study would reconstruct criminal careers by interviewing an unbiased random sample of completely honest offenders with perfect memories. Because such an approach is not feasible, we used DNA traces as a second-best alternative. But the use of DNA traces implies considerable selections in various phases of the investigative process. Before DNA traces are found at a crime scene, the crime has to be reported to the police, the police have to respond, and the crime and the crime scene need to be suitable for the collection of DNA material. Not every type of crime leaves DNA traces behind. Contact between the offender and the victim or the target is necessary, or the offender has to leave something behind that contains his DNA (e.g. clothes or cigarette butts). Crime scenes at which no DNA traces were found are missed when using DNA traces to study serial clearance. This selectivity of the data could have influenced our measurements. An offender who seems versatile on the basis of our data could in reality be specialized in a crime type that is not suitable for DNA research.

Although little is known about how the police handles recovering DNA traces from a crime scene, we do know that by reconstructing what happened and using information that is already known about the crime, the police will try to discover which DNA is offender DNA (Meulenbroek, 2008), and they will eliminate the victim or witnesses as donors of the DNA. Another important point when using DNA traces for criminological research is the reliability of a match between crime scene traces and between a crime scene trace and DNA retrieved from a person. A DNA profile is extremely rare, but it is impossible

DOI: 10.1002/jip

to know with 100% certainty whether a DNA profile is unique. It is not possible to find out if in the world population, nobody has the same DNA profile as somebody else. Therefore, there is always the possibility that a match between two traces is a coincidence. The Netherlands Forensic Institute calculated, however, that the probability that a match between two traces being a coincidence (when using a complete DNA profile) is always smaller than one in one billion (Meulenbroek, 2008). When considering this probability of a match being a coincidence, one should keep in mind that the same error margin has very different implications in research or in court. In court, an error margin of 1% may be unacceptable, whereas in social sciences, an error margin of 1% is negligible.

Although there are a number of limitations to be considered when DNA traces are used for criminological research, they also have one major advantage: they give us the opportunity to study the characteristics of crime series of offenders who were never arrested and compare these with characteristics of crime series of offenders who were arrested. By doing so, we have the opportunity to study which characteristics of these series influence the probability that the offender will be arrested. As shown in this paper, characteristics of crime series do have an influence on the probability that serial offenders will be arrested. Without the DNA data, we would not have been able to conduct this type of study and draw this conclusion.

Limitations of the study

Besides the remarks on using DNA traces for serial clearance research, another limitation of this study needs to be mentioned. There is a possibility that confounding variables have an influence on the results. These variables might have an influence on the probability of arrest, instead of the characteristics that are described in this study. We want to mention two of these possible confounders.

First, we could hypothesise that a versatile offender may be an opportunistic or chaotic offender. The result that versatile offenders have a greater probability of being arrested than specialized offenders could be a result of the offenders being chaotic in their offending and not so much a result of the versatility (and thus implicitly a lack of experience).

Second, interpersonal crimes such as violent offences and sex offences might be committed by offenders who are acquaintances of the victim. If the offender is known to the victim, the offender will probably be arrested quicker than if the offender is unknown to the victim. So, instead of seriousness (sex offences and violent offences are the two categories with the highest scores on the seriousness measure), it might be the fact that the offender is known to the victim that influences the probability of arrest. This possible confounding variable might also influence the specialization measures.

Future research

Interesting future research could be offender based. When was the offender arrested, or was there anything different about the crime he was arrested for compared with the crimes for which he was not caught? In other words, does the offender think that something in his behaviour might have influenced the probability of getting arrested?

Another possibility is to look at other characteristics of crime series and whether they have an influence on the probability of arrest. We could, for example, look at geographical aspects of criminal behaviour or at time aspects (e.g. the time of year or the time of day the crime was committed).

DOI: 10.1002/jip

A last possibility for interesting future research might be research into the practice of collecting DNA traces at crime scenes. Little is known about this. Of course, standards for DNA collection exist, but how it is carried out in reality remains unclear. Future research could look into how decisions about whether to collect DNA are made, when these decisions are made, and by whom. This type of research could give insight into the nature and extent of the selectivity of the DNA data.

REFERENCES

Addington L. A. (2006). Using national incident-based reporting system murder data to evaluate clearance predictors: A research note. *Homicide Studies*, *10*, 140–152.

Alderden, M. A., & Lavery, T. A. (2007). Predicting homicide clearances in Chicago: Investigating disparities in predictors across different types of homicide. *Homicide Studies*, (11), 115–132.

Blumstein, A., Cohen, J., & Nagin, D. (1978). *Deterrence and incapacitation: Estimating the effect of criminal sanction on crime rate*. Washington DC: National Academy of Science.

Clare, J. (2010). Examination of systematic variations in burglars' domain-specific perceptual and procedural skills. *Psychology, Crime & Law*, *17*(3), 199–214.

Cleves, M. A., Gould, W. W., Gutierrez, R. G., & Marchenko, Y. U. (2008). *An introduction to survival analysis using Stata*. Second edition. Texas: Stata Press.

Cox, D. R. (1972). Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society, Series B*, *34*, 187–220.

Davies, H. J. (2007). Understanding variations in murder clearance rates: The influence of the political environment. *Homicide Studies*, (11), 133–150.

Dodd, T., Nicholas, S., Povey, D., & Walker, A. (2004). *Crime in England and Wales 2003/2004 (no. 10/04)*. London: Home Office Statistical Bulletin.

Hazelwood, R. R., & Warren, J. I. (2003). Linkage analysis: Modus operandi, ritual, and signature in serial sexual crime. *Aggression and Violent Behavior*, (8), 587–598.

Heller, N. B., & McEwen, J. T. (1973). Applications of crime seriousness information in police departments. *Journal of Criminal Justice*, *1*, 241–253.

Jiao, A. Y. (2007). Explaining homicide clearance: An analysis of Chicago homicide data 1965–1995. *Criminal Justice Studies*, *20*(1), 3–14.

Keel, T. G., Jarvis, J. P., & Muirhead, Y. E. (2009). An exploratory analysis of factors affecting homicide investigations: Examining the dynamics of murder clearance rates. *Homicide Studies*, *13*, 50–68.

Kyvsgaard, B. (2003). *The criminal career: The Danish longitudinal study*. Cambridge: Cambridge University Press.

Litwin, K. J. (2004). A multilevel multivariate analysis of factors affecting homicide clearance. *Journal of Research in Crime and Delinquency*, (41), 327–351.

Litwin, K. J., & Xu, Y. (2007). The dynamic nature of homicide clearances: A multilevel model comparison of three time periods. *Homicide Studies*, (11), 94–114.

Liu, J., Francis, B., & Soothill, K. (2010). A longitudinal study of escalation in crime seriousness. *Journal of Quantitative Criminology*, *26*, 175–196.

Mazerolle, P., Brame, R., Paternoster, R., Piquero, A. R., & Dean, C. (2000). Onset age, persistence, and offending versatility: Comparisons across gender. *Criminology*, *38*(4), 1143–1172.

Meulenbroek, A. J. (2008). Research of biological traces and DNA-research. In A. P. A. Broeders & E. R. Muller (Eds.), *Forensic science*. Deventer: Kluwer.

Miethe, T. D., Olson, J., & Mitchell, O. (2006). Specialization and persistence in the arrest histories of sex offenders: A comparative analysis of alternative measures and offense types. *Journal of Research in Crime and Delinquency*, *43*(3), 204–229.

Nagin, D. (1998). Criminal deterrence research at the outset of the twenty-first century. *Crime and Justice*, (24), 1–42.

Netherlands Forensic Institute (2010). *Annual report 2009 of the Dutch DNA database for criminal cases*. The Hague: NFI.

Paré, P. P., Felson, R. B., & Ouimet, M. (2007). Community variation in crime clearance: A multilevel analysis with comments on assessing police performance. *Journal of Quantitative Criminology*, (22), 243–258.

DOI: 10.1002/jip

Ramchand, R., MacDonald, J. M., Haviland, A., & Morral, A. R. (2009). A developmental approach for measuring the severity of crimes. *Journal of Quantitative Criminology*, *25*(2), 129–153.

Regoeczi, W. C., Jarvis, J. P., & Riedel, M. (2008). Clearing murders: Is it about time? *Journal of Research in Crime and Delinquency*, *45*(2), 142–162.

Riedel, M., & Jarvis, J. P. (1999). The decline of arrest clearances for criminal homicide: Causes, correlates, and third parties. *Criminal Justice Policy Review*, *9*(3), 279–306.

Roberts, A. (2007). Predictors of homicide clearance by arrest. An event history analysis of NIBRS incidents. *Homicide Studies*, *11*, 82–93.

Roberts, A. (2008). The influences of incident and contextual characteristics on crime clearance of nonlethal violence: A multilevel event history analysis. *Journal of Criminal Justice*, *36*, 61–71.

Schneider, P. M., & Martin, P. D. (2001). Criminal DNA databases: The European situation. *Forensic Science International*, *119*, 232–238.

Sullivan, C. J., McGloin, J. M., Pratt, T. C., & Piquero, A. R. (2006). Rethinking the "norm" of offender generality: Investigating specialization in the short-term. *Criminology*, *44*(1), 199–233.

Tonkin, M., Grant, T., & Bond, J. W. (2008). To link or not to link: A test of the case linkage principles using serial car theft data. *Journal of Investigative Psychology and Offender Profiling* *5*, 59–77.

Topalli, V. (2005). Criminal expertise and offender decision-making: An experimental analysis of how offenders and non-offenders differentially perceive social stimuli. *British Journal of Criminology*, *45*, 269–295.

Woodhams, J., & Toye, K. (2007). An empirical test of the assumptions of case linkage and offender profiling with serial commercial robberies. *Psychology, Public Policy, and Law*, *13*(1), 59–85.

DOI: 10.1002/jip