WHAT WORKS CRIME REDUCTION SYSTEMATIC REVIEW SERIES

No. 13 A SYSTEMATIC REVIEW OF THE EFFECTIVENESS OF THE ELECTRONIC MONITORING OF OFFENDERS

Review Protocol

Jyoti Belur, Amy Thornton, Lisa Tompson, Aiden Sidebottom, Nick Tilley, Kate Bowers and Shane D. Johnson

UCL Department of Security and Crime Science

Contact address:

Jyoti Belur
UCL Department of Security
and Crime Science
University College London
35 Tavistock Square
London
WC1H 9EZ
England
j.belur@ucl.ac.uk

UCL Jill Dando Institute of
Security and Crime Science

what Works
Crime Reduction

E·S·R·C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL

College of
Policing

COPYRIGHT

Table of Contents

1. ABSTRACT

This protocol provides the background to, and reasons for, undertaking a systematic review of the evidence on the effectiveness of electronic monitoring (EM) of offenders. It outlines the questions that the review intends to answer and the methods through which relevant studies will be identified, appraised and synthesised. Guided by the recently established EMMIE framework (Johnson, Tilley and Bowers 2015), the review has three broad aims: 1) to establish whether the EM of offenders has been found to be effective as a method of reducing crime; 2) to investigate how, in what form(s), and under what conditions the EM of offenders has been found to be effective; ineffective; and/or to produce unintended negative effects; and 3) to collate information on the implementation and costs of electronically monitoring offenders. The review will combine meta-analysis, typically associated with Cochrane and Campbell Collaboration systematic reviews (see Sackett et al. 1996; Petrosino et al. 2001), and realist synthesis (Pawson and Tilley 1997; Pawson 2006) in the tradition of other reviews conducted under the purview of the *What Works Centre for Crime Reduction*, hosted by the UK College of Policing (see, for example, Sidebottom et al 2015).

2. BACKGROUND

The move towards promoting EM technology is rooted in the penal crisis of the 1970's and 1980s when prison overcrowding and the rapidly escalating cost of incarcerating increasing numbers of offenders was recognized as a problem in the USA and the UK (DeMichele and Payne 2009, Paterson 2007). EM was designed, in part, to prevent the need for offenders to remain in custody, be that in pre-trial detention or through fulfilling their entire sentence length. This allows governments to reduce costs by providing cheaper alternatives to custody (Garland, 2002).

Electronic monitoring, also commonly known as tagging, is used as a component of sentencing for juveniles and adults to enable release of offenders who might otherwise be imprisoned as part of Curfew Orders, or allow early release of prisoners on Home Detention Curfews (House of Commons Committee of Public Accounts 2006). Curfew orders vary, but generally involve requiring offenders to be present at a curfew address for a fixed number of hours per day for the duration of a pre-determined period. Tagging, in the UK, is accomplished by means of a device attached to the offender (usually ankle) and the installation of monitoring equipment at the curfew address (House of Commons Committee of Public Accounts 2006). The tag itself

does not prevent reoffending but merely monitors the whereabouts of offenders. The decision therefore to recommend monitoring usually depends on a competent authority's assessment of offender suitability for EM.

EM technology has developed considerably in recent decades. Initial systems in the 1980's involved bracelets on the ankle or wrist of the offender and a monitoring device fitted in the offender's home location which sent signals, alerting the authorities if the offender strayed too far from their home (Renzema and Mayo-Wilson, 2005). The late 1990's saw the development of GPS monitoring systems logging offenders' movements, which would allow the police to know their whereabouts at any time of the day or night. EM systems may be implemented alone, or as part of a suite of interventions including intensive supervision by parole or probation officers alongside drug treatment (Renzema and Mayo-Wilson 2005). All EM monitoring schemes have some private sector involvement, ranging from development and implementation of technology to partial or full control of administration of curfew orders. For example in the UK three companies - G4S, Serco and Sodexo - hold all prison contracts and the national contract for EM is held by Capita (Paterson 2007, Whitehead et al 2013).

A large number of jurisdictions within the western world utilise EM technology, including England and Wales (Mair 2005, Paterson 2007), Scotland (Smith 2001), New Zealand (Gibbs and King 2003), Australia (Black and Smith 2003), Sweden (Marklund and Holmberg 2009), wider Europe (Haverkamp, Mayer and Levy 2004) and the USA (Austin and Hardyman 1991, Baumer and Mendelsohn 1991, Jones and Ross 1997, Lipner 1993). The reach of monitoring has extended to cover those on bail, adult and juvenile offenders, terror suspects, those suspected of breaching immigration laws and potentially in the future those refusing to pay child support (Paterson 2007).

In England and Wales, the Home Office have conducted a number of studies assessing the efficacy of EM programmes (National Audit Office 2006, Dodgson et al 2001, Mair 2005, Mair and Nee 1990, Mair and Mortimer 1996, Shute 2007, Sugg, Moore and Howard 1991), and remain interested in the cost-saving abilities of such devices, with one study showing that a 90 day curfew period with a tagged offender is around £5,300 cheaper than a custodial sentence of the same length (National Audit Office 2006). The National Audit Office study also concluded that the efficacy of EM depends on the quality of the performance of contractors' responsible for installing and maintaining equipment and monitoring offenders. The study also

identified certain implementation problems with EM including delays in conducting assessments for suitability of offenders for EM and also delayed criminal justice response to reported breaches. These findings were echoed in other studies. The Home Detention Curfew Scheme in Great Britain has seen EM flourish, with 137,000 people on the scheme between 1999 and 2006 at a cost of £342 million with monitoring companies charging an average daily cost of £42 for tracking a person passively (Shapps 2006, Shute 2007). In the USA, EM devices have previously cost as little as $3 a day to maintain (Glaser and Watts 1992). Since the 1990s, the numbers of people subject to EM has reduced in the USA, although around 125,000 people are subject to EM based curfew orders and around 8.000 to 9,000 people are being tracked by GPS systems (Stacey 2006).

A range of different crime types can be addressed by EM, with domestic violence, violent crime, and sex offences all being studied (Erez et al 2012, Finn and Muirhead-Steves 2002, Payne and DeMichele 2011). In the latter case, sex offenders with tags may be informed that they are not allowed within a certain distance of schools, playgrounds and other areas in which children commonly congregate. EM can also be implemented at different points within the criminal justice system, including as an alternative to pre-trial detention, as a stand-alone community sentence, as part of a restraining order in the case of domestic violence perpetrators, or after a sentence has been served (as early and temporary release, or parole). It could also serve to assist offender re-entry into the community and there is some evidence that drug and/ or sex offenders on EM are more likely to complete treatment than other offenders (DeMichele and Payne 2009, Crowe et al 2002).

EM is not without its critics, with ethical issues being raised about the intrusiveness of the technology, the impact upon the offender and their family, and how it may adversely affect minority ethnic groups as a result of shame associated with EM (Payne, DeMichele and Okufo 2009). Technological issues such as the ability to remove tags, false alarms and inappropriate implementation have all been documented (DeMichele and Payne 2009, House of Commons Committee on Public Affairs 2006). The commercialisation of EM due to government subcontracting has also been contentious (Paterson 2007). While EM has the potential to enhance supervision of offenders in the community, its efficacy is only as good as the technology. Moreover, increased supervision demands could cause supervisory officers undue stress and lead to other negative consequences (DeMichele and Payne 2009). Attitudes towards EM are not, therefore, exclusively positive.

The authors are aware of one past systematic review of the evidence of EM (Renzema and Mayo-Wilson 2005). That review found only three studies that met its stringent inclusion criteria, with the combined conclusion of these studies finding no impact upon recidivism rates after one to three years. There are several reasons why we believe an updated and extended review of the effectiveness of offender monitoring is warranted. First, the review of Renzema and Mayo-Wilson (2005) is now over 10 years old. Since then, EM technology has improved (DeMichele and Payne 2009). Second, Renzema and Mayo-Wilson's review was limited to primary studies that a) were quasi-experimental in design and b) focused on moderate to high-risk offenders. This meant that only three studies were eventually included by the authors. Further primary studies evaluating the effectiveness of EM have been published since then (for example, National Audit Office 2006, Shute 2007, Bales 2010, Bulman 2013). Moreover, primary studies using different methods may be included to widen the evidence base. Given the importance of EM worldwide and the rapid improvement in monitoring technology over the past few years, it is important to ensure that the best available evidence is synthesised and made available to practitioners in order to accurately inform their decision-making in this area. Third, the existing review provides little information about the mechanisms through which EM might prevent crime, the contexts in which EM works best, and any implementation and economic considerations. EM has been dogged by technological and implementation failures, on the part of private contractors in monitoring, as well as criminal justice agencies in responding to breaches, thus reducing the immediacy of sanctions and affecting its deterrent effect on offenders (National Audit Office 2006, House of Commons Committee of Public Accounts 2006, Shute 2007). This review will synthesise information on mechanisms, costs, contexts and implementation issues to aid practitioners and policy makers in this area.

THEORETICAL FOUNDATION AND FOCUS

EM is theorised to work by reducing an offender's opportunity to commit crimes (Mackenzie 2006). It is used as an additional tool to other risk management techniques such as hostels, exclusion zones, curfews, drug orders and public protection multi agency oversight (Shute 2007). If the offender is under some kind of temporal curfew, they are no longer at liberty to leave a designated place (usually a home location) at certain times. They are thereby deterred from committing offences other than in that location during those hours. It may be that they are banned from certain physical areas (such as city centres, schools or playgrounds), from

associating with certain people (such as other known offenders or victims), or being within a certain distance of a known crime hotspot. These physical and temporal restrictions limit the offenders' capacity to commit crimes. By placing offenders under surveillance through technology such as GPS, which can track their location at all times, offenders either know or perceive that they can be traced and caught if they break the conditions of their monitoring agreement. This should, theoretically, increase their perceived risk of being caught and thus deter them from committing crime. At the same time restricting their access to crime hotspots, at particular times, or their access to known offenders, reduces the opportunity to commit offences. Additionally, by serving as a reminder not to offend, EM has been theorised to provide a 'psychological reinforcement' to deter offenders from committing crimes (Shute 2007). Satellite tracking can also be used to enhance supervision of offenders on curfew, gather intelligence about offenders and use it to link them to offences or eliminate them from suspicion thus aiding in detecting crime (Shute 2007).

There are various types of technology which can be considered EM (DeMichele and Payne 2009):

- Continuous signalling house arrest device. These are placed on the offender and emit a signal every minute or more, transmitting the offender's location to another device within the home to confirm their proximity.
- Mobile monitoring devices. These can be used by law enforcement officials who can then detect the location of an offender wearing a tag within a certain proximity. Officers can drive by offenders' homes and ensure they are present.
- Location tracking devices (GPS). These usually comprise portable tracking devices that the offender must carry with them in combination with a tag worn by the offender. They are able to constantly record the location of the offender in almost real time.
- Programmed contact systems (voice verification, video verification, device verification). These are automated calling systems designed to check if an offender is in a given location at a certain time. The offender may need to answer the telephone for voice recognition, be visible for video verification, or be close to device verification software.
- Remote alcohol/drugs detection devices (periodical or continuous). These continuously or periodically test for the amount of alcohol or drugs in an offender's system as the misuse of such substances may be part of the monitoring conditions.

- Victim protection devices. These can be located in a victim's home to alert them if an offender is within a certain proximity (most often used for domestic violence cases).

This review will focus exclusively on studies of EM that include some sort of tagging device worn by an offender, and hence exclude studies relating to monitoring which is solely conducted by using CCTV or via telephone. This is because the focus of the review is on monitoring as opposed to general surveillance. Tags are used to monitor specific individuals (offenders or suspects) but CCTV is used more generally as an indiscriminate surveillance measure. The studies must also focus on EM of offenders as opposed to goods or places and must have a crime prevention outcome. For example studies focussing upon RFID tracking systems of goods will not be included in this review.

As this review is interested in the historical development of EM devices, we will include the technological devices used in a moderator analysis (to test for differential effects of different device types) if enough information is available for this. The devices of interest would include verifier wristlets, anklets and bracelets. This review is concerned with physical tags on offenders, rather than telephone voice recognition systems, which do not require a tag but only for an offender to answer the phone at a certain time. Those systems that use telephone voice recognition to ascertain an offender's whereabouts if their worn tag has indicated that they have broken their curfew are of interest.

3. THE APPROACH OF THIS SYSTEMATIC REVIEW

Johnson et al. (2015) proposed EMMIE as a coding system to assess the quality of evidence generated by systematic reviews. It identifies five strands of evidence relevant to the decision making of crime prevention policy-makers and practitioners. The first 'E' of EMMIE denotes the size of the 'effect' of a given policy, programme, practice or intervention. The initial 'M' refers to 'mechanism' – what it is about a policy, programme, practice or intervention that brings about its effect. The second 'M' refers to 'moderator' (or what realists refer to as 'context'). This describes the conditions that need to be in place if the policy, programme, practice or intervention is to activate the mechanisms necessary to produce given effects. The 'I' refers to 'implementation'. Decision-makers need to know whether and how a policy, practice, programme or intervention can be put in place. Finally, the second 'E' refers to 'Economics' – what the intervention will cost in relation to outputs, outcomes or benefits.

In addition to appraising the evidence of existing systematic reviews, Johnson, Bowers and Tilley (2015) suggested that EMMIE might usefully guide the conduct of new systematic reviews. The review by Sidebottom et al. (2015) of the effectiveness of alley gating as a method of reducing crime was the first to do so, combining two distinct and hitherto rarely integrated approaches to systematic review: statistical meta-analysis and realist synthesis. Each asks different yet complementary review questions. The meta-analysis will therefore focus on whether EM of offenders has been found effective as a method of reducing crime. The realist component of our review is concerned with finding out how, in what form(s), for what crime types, for whom, and under what conditions EM has been found to be effective, ineffective and to produce unintended negative effects.

More specifically, the review aims, where possible, to generate evidence regarding:

a) The mechanisms involved with EM for the purposes of crime reduction for offenders in particular spaces at particular times (for example increased risk of being caught in breach, reduced opportunity to commit crimes in vulnerable places or at particular times);

b) The causal mechanisms through which EM techniques have been found (positively or negatively) to affect crime, and whether these mechanisms vary substantively by crime type and/or specific sub-groups intended to receive the intervention (for example restricting access to physical spaces at particular times thus making offending more difficult or more risky in real or perceived terms);

c) The causal mechanisms through which EM techniques have led to the achievement (or otherwise) of other secondary outcomes (such as treatment completion for drug or sex offenders on probation or parole, or completing conditions of parole)

d) The conditions found to be (more or less) conducive to the activation of causal mechanisms identified in (b) (for example prompt law enforcement response to reported breaches of home detention rules, functionality of the technology);

e) The conditions found to support or hinder the implementation of various kinds of monitoring devices and, for example, conditions necessary for home detention access-blocking interventions of the kind identified in (a) (for example, available funding and key stakeholder agreement for the installation and maintenance of EM device);

f) The degree to which EM measures taken are cost-effective/otherwise beneficial as well as effective at reducing crime;

g) The patterns of (positive and negative) crime-related outcomes produced by (a)-(e) (for example overall positive impact on offender lifestyles)

4. REVIEW METHODS

This section describes the methods to be used in this systematic review. It begins with an overview of the strategy to be taken, in particular how the proposed methods speak to both meta-analysis and realist synthesis. This is followed by a description of our inclusion criteria, strategy for identifying studies, search terms, data extraction and management processes, and the statistical analyses to be performed in the meta-analysis.

The approach taken here is consistent with that reported in Sidebottom et al. (2015). Briefly, using EMMIE (Johnson et al. 2015) as a guiding framework for the types of evidence necessary to inform improved decision-making, we combine two distinct approaches to systematic review: statistical meta-analysis and realist synthesis. Each approach is directed toward a different yet complementary set of review questions. Our meta-analysis is primarily concerned with systematically appraising the evidence to determine 1) whether EM 'worked' to reduce reoffending 2) what factors were found to moderate effectiveness and 3) whether EM was a cost-beneficial theft prevention measure. Our realist synthesis is primarily concerned with systematically appraising the evidence to uncover 1) how EM works to reduce reoffending or produce significant side-effects (what realists call the 'mechanisms'), 2) the conditions necessary for electronic tags to activate mechanisms that reduce reoffending (what realists call the 'context') and 3) the outcome patterns observed if EM operates as expected. These two approaches to review should not be considered independent, however. It is our hope that the findings of each branch of our review can inform and be informed by the other. For example,

emerging context-mechanism-outcome pattern conjectures generated by the realist review may be empirically tested with the data available for meta-analysis if the evidence collected in the primary studies is adequate. It is our contention that pursuing evidence that speaks to the range of review questions raised by EMMIE will maximise the value and usefulness of the outcomes of this review for industry, researchers, practitioners and policy makers.

Brief overview of review process

As illustrated in Figure 1 below, our review is expected to unfold as follows. Relevant databases (including the grey literature, see pp.19-20) will be searched using pre-determined search terms (pp. 21). Identified studies will initially be screened through reading the title and abstract to remove those that are unsuitable based on our inclusion criteria (pp. 18-19). Crucially, as explained below, the inclusion criteria for the realist synthesis differ from those for the meta-analytical branch of this review. For example, to be included in the meta-analysis, a study must report a quantitative crime-relevant outcome measure (following the implementation of EM). For the realist synthesis broader inclusion criteria will be used. Studies that meet the respective inclusion criterion will then be coded to extract relevant information.
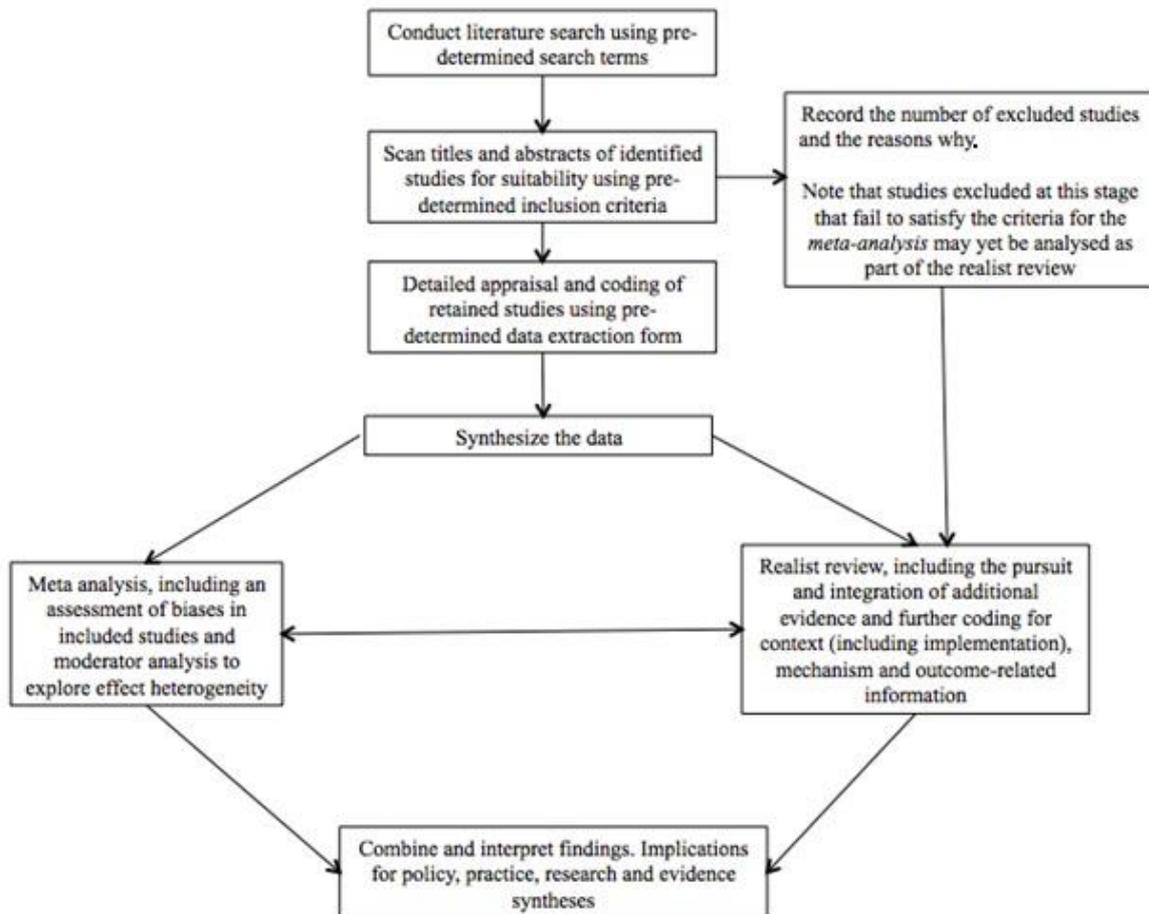
Figure 1: Flowchart of the Process to be followed in our Systematic Review of EM of Offenders as a Method to Reduce Crime

In simple terms, the two branches of our review can be thought of as centring on different - albeit related - questions. The meta-analysis branch (to the left of Figure 1) is primarily concerned with "what worked?" in the context of evidence on the effectiveness of EM as a method of crime reduction. The realist branch (to the right of Figure 1) draws on the same material as the meta-analysis, as well as additional sources of evidence considered relevant, to answer the broader question "what works, how, for whom, and under what circumstances?"

Once identified, studies that have been screened in according to the inclusion criterion will be drawn on to follow two analytical paths, co-ordinated by two separate (but interacting) research teams. One team will quantitatively synthesise the studies with meta-analytic methods to generate an overall effect size. Where possible, moderator analysis will be conducted to explore

any heterogeneity in effect sizes observed across studies and differences between sub-groups. An assessment of potential biases (such as publication bias) in the included studies will also be carried out. In parallel, the same studies (plus additional ones) will be analysed from a realist perspective by the other team. This will involve a detailed reading of the selected studies with a view to eliciting working theories as to how EM is expected to effectively reduce crime in the different offenders who have been tagged and to produce other unintended consequences. There are several features to the realist analysis that warrant mention. It will involve some extra coding to that conducted in the meta-analysis, pertaining to:

1) the causal mechanisms through which the EM measures were deemed to work, and where possible the coding of intermediate outcomes that would be expected if a particular mechanism were at play;
2) implementation-relevant information, and
3) outcome patterns attributable to the EM measures that extend beyond net increases or decreases in crime. Such outcome patterns may include, for example, increase in offending behaviour outside of monitored hours or change in modus operandi or location of offending behaviour outside of restricted areas.

While nuanced coding categories will be generated inductively to help summarise emerging themes, the coding will also be binary, indicating whether studies do or do not report such items. We consider this to be an important element of any systematic review: informing not just what we already know but what we need to know. More specifically, we expect this information to usefully inform the development of reporting guidelines for criminological studies, should it emerge that a sizable proportion of relevant studies make no mention of these issues. Presently, reporting guidelines for criminological research have been mooted, drawing on established checklists in the health sciences, but not yet formally developed (see Sidebottom and Tilley, 2012; Perry, Weisburd and Hewitt, 2010). Growth in the use of systematic review methods in criminology suggests that the development of reporting guidelines is timely. The research conducted as part of the "What Works?" project is expected to advance this agendum.

In the spirit of realist synthesis, should evidence or information gaps emerge in the process of analysing the identified studies, this will initiate a further purposive search for relevant evidence. This pursuit of pertinent evidence hitherto unidentified in the search process is intended to produce a richer understanding of *the conditions under which* EM reduces crime

(regardless of the type of evidence, quantitative or qualitative, or research design employed), and to help build and refine underlying programme theories. To illustrate what is meant by sourcing additional relevant evidence, consider a hypothetical scenario in which a study described the time delays in response by law enforcement to breaches of curfew, or the delays and wastage caused by inconsistency in initial assessment, or delays in getting necessary paperwork to enable release on tagging leading to offenders being released after their eligibility dates. This might usefully inform future strategies to control costs that seek to implement similar crime reduction measures. In realist terms, this would relate to the conditions necessary to activate the causal mechanism of increasing risk and reducing opportunity, believed to be responsible for producing the outcome of interest. Further, studies identified in the initial searches may be silent on issues of implementation – something that would be important to practitioners seeking to implement EM. Other studies, which fail to meet the inclusion criteria for meta-analysis, might provide useful guidance on such matters. The realist arm of the review would aim to identify such studies.

We recognise that additional searching of the sort described above may be very demanding, and we will be limited by the time and resources available. Our hoped-for goal, following Pawson (2006), is to marshal sufficient evidence to build, refine and test theories judged to underlie the intervention under review (here EM of offenders) until the point of 'theoretical saturation', whereby consensus is reached that sufficient evidence has been gathered to answer the review question(s). We also recognise that this type of review is not linear, but describes an iterative procedure, with feedback loops potentially taking researchers back to literature that they considered earlier in the process. This flexible searching for evidence that can inform theory lies at the heart of realist approaches.

It must be stressed that we do not envisage these dual – realist and meta-analytic – processes as being independent. On the contrary, we expect the emerging findings from one branch of evidence synthesis to inform and be informed by those from the other. One example where we see this cross-over as being particularly likely is with respect to the testing of explanations for any observed heterogeneity in effect sizes across studies. Briefly, reviews into the effectiveness of crime prevention interventions invariably uncover substantial differences between studies, types of interventions, time periods, population groups and so on. Heterogeneity in effect sizes is the norm. Moderator analysis has therefore come to be standard practice in systematic

reviews (most notably meta-analyses) in an attempt to locate the source of and reasons for sub-group differences. This typically involves partitioning the accumulated data (and hence studies) into meaningful subsets so as to calculate mean effect sizes.

If the information from primary studies is available, we aim to analyse the effect of a number of potential moderators which may have an impact upon the size of the effect, if any, that EM is found to have upon reoffending. These include:

- Adult versus juvenile offenders
- Male versus female offenders
- Offence type – violent crimes, sexual offences, drug/alcohol-related offences
- Offender criminal history – first time offenders versus repeat offenders/career criminals
- Geographical location – different countries (e.g. USA vs UK) and states/counties within USA and UK if possible
- Isolated intervention or part of a suite – just tagging or other things like intensive supervision (frequent meetings) and drugs monitoring
- Technology type – GPS versus other devices
- Stage of criminal process – pre-trial detention, early release, parole, stand-alone
- Sanctions for violating monitoring – seriousness of sanction may affect behaviour
- Study date – by decade (this may interact with technology type and location)
- Study design – RCT versus quasi-experimental studies
- Study sample size – larger versus smaller participant numbers

Presently, the choice of variable on which moderator analyses are conducted tends to be determined by tradition and convenience as opposed to generative causal mechanisms judged to be responsible for the outcome patterns observed (or not). For example, regardless of the question under review, moderator analysis typically sifts the data to determine any significant differences by, say, study location, intervention type and study design. These findings are clearly of scientific interest and may reveal interesting results, although the risks of *post hoc* atheoretical partitioning have been long recognised in the meta-analyses of medical trials (e.g. Peto et al. 1995; Rothwell, 2005a; 2005b). Moreover, such analysis may be less relevant to *evidence consumers*, most notably policy-makers and practitioners, interested in the ways in which the findings may be applicable to the practicalities of preventing crime in their particular

setting. This is again an issue that has been stressed in clinicians' discussions of the use of meta-analysis of trials in their medical practice (*e.g.* Horowitz, 1987, 1995; Feinstein 1998; Davidoff 2009). As alluded to above, one aim of the realist branch of our review is to uncover and test working theories as to how and under what conditions EM will reliably reduce (or fail to reduce) reoffending. To this end, it is our hope that a realist perspective might inform a theory-driven and also more practically meaningful moderator analysis. In particular, it aims to generate theoretically warranted analysis of sub-groups.

Where *post hoc* moderators are introduced in the meta-analysis in light of evidence of heterogeneity in outcomes uncovered across multiple studies, the realist branch will try to explain the findings and also to search out further data that might be drawn on to support or contradict suggestive explanations by reference to the distinct expected outcome signatures (Johnson, Birks, McLaughlin, Bowers and Pease, 2007; Eck and Madensen 2009; Pawson and Tilley 1997). As with any systematic review, the extent to which meaningful moderator analyses can be conducted will be determined by the number of studies meeting our inclusion criteria and the data available from them.

Criteria for considering studies for this review

In selecting studies for this review we will use the following inclusion criteria:

a) *The study must have reported an explicit goal of reducing crime (i.e. reoffending) through the use of EM.* We will include studies on EM implemented by any stakeholder: law enforcement, criminal justice agencies, public sector and private entities, etc. We will also consider studies that report the effects of EM implemented in isolation and as part of a wider package of crime reduction interventions. The study must also focus on EM of offenders as opposed to goods or places.

To be included in our meta-analysis, a study has to satisfy point a) above and:

b) *A study has to report at least one quantitative crime outcome measure.* Outcome data could comprise official measures (e.g. reconviction or arrest data) or unofficial measures (self-reported levels of offending and/or victimization; breaches of curfew

conditions). Studies that only report non-crime-related outcome measures (such as perceptions of effectiveness and technological or ethical issues) will not be included.

c) *A study has to report original research findings.* Systematic reviews will not be included. The quantitative findings for any single study will only be incorporated once, even if reported in multiple publications. Where this is the case, the study reporting the most detailed information will be included or, where necessary, any dependency in the data will be dealt with appropriately.

d) *A study has to employ a research design that permits the computation of a reliable effect size (i.e. an experimental or quasi-experimental evaluation design with control group or a suitable single study interrupted time series design).*

The shortage of experimental and quasi-experimental studies in criminology is well recognised, particularly for evaluations of place-based situational interventions (Eck, 2006; Guerette, 2009). This can result in only a small number of studies being eligible for meta-analysis. Because of this, while following the above criteria we will also consider studies that measure the impact of EM using simple before and after designs and no control group. Where such studies are used in the quantitative analysis, this will be clearly indicated, as will the familiar concerns regarding the internal validity of this approach.

Items *b*, *c* and *d* did not form part of the inclusion criteria for the realist synthesis. To be included in our realist synthesis of EM, studies had to satisfy point *a)* above - report an explicit goal of reducing crime through the use of EM - *and at least one* of the items below:

e) Report substantive information relating to crime-related causal mechanisms activated by EM
f) Report substantive information relating to the conditions needed for EM to activate crime-related causal mechanisms
g) Report substantive theoretical content concerning EM and crime-related outcomes

For items e to g, *substantive* is understood as any information that contributes to the theory, design, implementation and evaluation of EM to reduce crime.

More generally:

    h) Published and unpublished studies will be included

    i) There will be no restriction by date of study

    j) Studies must be available in the English language. Available resources limit our ability to search and translate non-English EM-relevant studies

Identifying studies: databases and information sources

Relevant studies will be identified using the following search methods:

1) A keyword search (see p 21) of relevant online abstract databases, including grey literature and dissertation databases (see p 19-20)
2) Forward and backward citation searches of all studies that meet our meta-analysis inclusion criteria.

The finalised list will be checked by recognised experts in the field and an information specialist at Rutgers University (Phyllis Schultz) who has extensive experience in retrieving articles, particularly from the grey literature in policing and crime prevention.

Databases to be searched are:

1. ASSIA (Applied Social Sciences Index and Abstracts)
2. Criminal Justice Abstracts
3. Criminal Justice Periodicals
4. ERIC (Education Resources Information Centre)
5. IBSS (International Bibliography of Social Sciences)
6. NCJRS (National Criminal Justice Reference Service)
7. ProQuest theses and dissertations
8. PsycINFO
9. PsycEXTRA
10. SCOPUS
11. Social Policy and Practice
12. Sociological Abstracts

13. Web of Science

14. CINCH

In collaboration with Phyllis Schultz, an information specialist and librarian at Rutgers University (US), we will also search the publications of the following government, research and professional agencies:

1) Centre for Problem-Oriented Policing (Tilley Award and Goldstein Award winners)

2) Institute for Law and Justice

3) Vera Institute for Justice (policing publications)

4) Rand Corporation (public safety publications)

5) Police Foundation

6) Police Executive Research Forum

7) The Campbell Collaboration reviews and protocols

8) Urban Institute

9) European Crime Prevention Network

10) Swedish National Council for Crime Prevention

11) UK Home Office

12) UK Ministry of Justice

13) UK College of Policing (Polka)

14) Australian Institute of Criminology

15) Swedish Police Service

16) Norwegian Ministry of Justice

17) Canadian Police College

18) Finnish Police (Polsi)

19) Danish National Police (Politi)

20) The Netherlands Police (Politie)

21) New Zealand Police

22) US National Institute of Justice

Finally the following resources will also be utilised:

1) Google

2) Google Scholar

3) Academic Search Premier (EBSC)

4) ProQuest Sociology

5) Rutgers Criminal Justice Grey Literature Database

6) OSCE Polis Digital Library

Full text versions of identified studies will be obtained through one of the following means (in order of preference):

a) Electronic copies via University College London e-journals service (UCL; as well as other electronic works accessible through other universities as part of a consortium, *e.g*. University of London Senate House Library).

b) Electronic copies of studies available from elsewhere on the internet.

c) Paper copies at UCL (as well as other electronic works accessible through other universities as part of a consortium, *e.g*. M25 consortium).

d) Electronic/paper copies requested through UCL's Inter Library Loan (ILL) system, which sources most materials from the British Library.

e) Electronic copies at the Gottfredson Library at the Rutgers University School of Criminal Justice (US).

f) Electronic/paper copies requested from the authors themselves.

g) The UK College of Policing library.

Should any of the full text versions of the works collated contain insufficient information to determine their eligibility for inclusion according to our coding strategy (described below), where practicable the corresponding author will be contacted in an attempt to retrieve this information.

Search strategy/terms

Our search terms will be based on two concepts: EM or variations thereof and crime outcomes.

i. Thus we will use the following search terms for the intervention:
"electronic monitor*" OR tag* OR curfew* OR "random calling" OR "verifier anklet" OR "verifier wristlet" OR "verifier bracelet" OR ((house OR home) W/1 (arrest OR detention OR confinement OR incarceration))

ii. We will use Boolean operators to combine it with the following search terms for the crime outcome:

"crim* OR illegal* OR illicit* OR delinqu* OR offend* OR parole* OR probate* OR incarcerate* OR recidivism* OR convict* OR felon* OR misdemeano* OR bail*"

After a pilot test, it was decided that the following terms would not be used:

secur*, suspect*, prohibit*, GPS, RF, satellite monitoring system

These terms were found to have widened the search and resulted in too many studies returned which were not related to crime or crime prevention. It was also discovered that those relevant studies which did use the terms GPS, RF and satellite monitoring system were still being discovered with the other search EM terms above.

Data extraction and management

As indicated in Figure 1, the first level of screening involves the review team examining the title and abstract of those studies returned following the initial electronic and bibliographic searches. All references will first be uploaded using EPPI 4 reviewer software, a web-based software program developed by the Social Science Research Unit at the Institute of Education, University of London, to manage and analyse data generated from systematic reviews (http://eppi.ioe.ac.uk/cms/Default.aspx?alias=eppi.ioe.ac.uk/cms/er4). Once uploaded, studies will be assessed using our inclusion criteria. Four possible outcomes are possible at this point: 1) studies are excluded from the review for not meeting our inclusion criteria, 2) studies are retained for definitely meeting the inclusion criteria, 3) studies are retained for possibly meeting the inclusion criteria and require reading of the full text to decide definitely and 4) studies are retained for follow-up as they sign-post another potentially relevant study (such as a book or article review). Rates of attrition and the number of studies in each of these four conditions will be reported. Any disagreements over whether a study should be included or excluded will be resolved by discussion between the research team. Inter-rater reliability will also be examined (see page 27).

The next stage involves screening the full text of, and extracting data from, retained studies. This will be done separately for studies that meet the inclusion criteria for meta-analysis and those studies that meet the inclusion criteria for realist synthesis. The following information, where available, will be coded from studies that meet our meta-analysis inclusion criteria:

1. Study details (title, year, author(s), author affiliation, publication, study location(s))

2. Nature (type) of EM intervention(s) put in place

3. Dosage (intensity) of intervention

4. Type(s) of crime(s) examined

5. Unit of analysis (people, cohorts, etc.)

6. Research design (RCT, quasi-experimental, pre-post test)

7. Description of comparison group, place or period

8. Sample (size and any notable features)

9. Statistical test(s) used

10. Outcome measure of interest and data source (curfew conditions breached, arrest, recall to prison etc.)

11. Effect sizes (where applicable and/or reported) will initially be coded as they are reported (but see below)

12. Other interventions implemented over the study period

13. Indirect effects (displacement and/or diffusion of crime control benefits)

14. Financial costs and benefits

15. Conclusions of the author(s)

As part of the realist stream, the full text of those studies that meet the realist synthesis inclusion criteria will be screened and information related to the topics below will be coded. As already described, absence of such information will also be logged.

1. Causal mechanism(s) judged to be responsible for the sought-after (observed) outcomes

2. Context of intervention: nature and description of the setting in which the EM measure was implemented. This can be considered from several perspectives. Context might refer to, in one instance, the method by which assessments of suitability for EM are conducted and the role and inputs of various agencies in the decision-making process. In another sense it might also refer to the type of offender chosen, the type of offences committed, the stage at which the decision to monitor electronically was taken and whether the tie-up between the private company monitoring offenders and law enforcement is well established to ensure prompt response to breaches of tagging conditions. Contextual factors associated with differential effectiveness might also relate to time of day or season. We expect to code this sort of information in a free-text format with emerging categories being developed over time.

3. The process of and conditions under which the EM measure was implemented, as well as the different stakeholders involved in their individual roles and responsibilities.

4. Intermediate outcome measures that might help to understand the mechanism(s) through which an interventions brings about its effects.

All information will be stored coded using an appropriate coding framework in EPPI 4.

5. STATISTICAL PROCEDURES FOR META-ANALYSIS

We expect studies to differ in methodological approach. Some will have simple post assessments, some pre and post assessments, and some pre and post assessments with at least one control group. We predict that fewer still will have established some equivalency between comparison and treatment groups, or will have used random assignment to minimise bias. For completeness, those studies that employ research designs that limit confident determinations of impact will be examined *prima facie* for their presence, but they will not be the focus of the meta-analysis.

To account for the varying levels of methodological rigor, studies will be grouped according to a hierarchy of evidence (most likely quasi-experimental designs versus RCTs) – which reflects the extent to which causal inferences will be sensible – and analysed separately. Estimates of effect size (ES) will be computed within groups and comparisons made between them (see below). The statistical element of the quantitative review will focus on studies which at least meet the following conditions: i) they presented raw counts or rates of crime, arrest or recidivism, or reported a standard measure of effect size and sampling variance that is suitable for inclusion in meta-analysis; and ii) the research design used in the evaluation comprised at least two groups: a treatment and a control. These may use random or non-random assignment to conditions (treatment and control), although very few are expected to have used randomization in the design.

It is expected that a range of different data and methods will be used across primary studies. Some of the issues anticipated will be the use of multiple treatment and control groups; the availability of time series data; the use of different outcomes; the use of different methods for calculating effect sizes; and the computation of effect sizes for different crime types and differences in the handling of intervention effect sizes. These differences will be identified during the coding stage and dealt with in the statistical analysis (see below).

Calculating effect sizes

To estimate the effectiveness of EM interventions overall, and for relevant sub-groups of studies, we will undertake a statistical meta-analysis. To enable synthesis, the individual effect sizes – which may be reported using different test statistics (*e.g.* odds ratios, mean difference scores, and so on) – will be converted to a common metric. To do this, we will standardise by converting to the type of effect size that is most frequently used across the primary evaluations. In other cases, we may find that outcomes are reported as F-ratios or as standardised differences in means statistics. Where possible, if the original authors have not calculated an effect size, but it is apparent that this would be possible using the available data, we will endeavour to calculate an effect size.

Having converted the effect sizes to a common metric, we will undertake a statistical meta-analysis. In line with contemporary research, to deal with any heterogeneity in the effect sizes, we will compute a mean effect size using a random effects model. As is standard practice, when combining effect sizes to compute an overall mean effect, individual metrics will be weighted using inverse variance weights. This will ensure that more reliable effect sizes are given more weight in the calculations. Along with the overall mean effect size, individual effect sizes will be presented, most likely using forest plots showing point estimates and the associated confidence intervals.

Dealing with dependency

It is likely that there will be cases where it is possible to generate more than one effect size from a single primary study. Reasons for this will vary and we could therefore deal with them according to the particular situation, as follows:

1. Data are presented for multiple treatment groups, each with independent matched controls. Where this occurs we will take the mean of the available effect sizes and use this as the overall outcome for the primary study.
2. Treatment groups are compared to more than one control group. In this case, one option is to compute two effect size measures for the study, one showing the worst case scenario and one showing the best. The overall mean effect size (computed across sites) could then be computed using data to show: a) the best case scenarios; and, b) the worst

case scenarios. Forest plots summarising the effect sizes will be produced for each scenario.

3. Reductions in crime in treatment groups compared to control groups using different pre and post time periods. We will deal with this as in (2) above.

One issue with these approaches is that data are lost or averaged. Therefore, we will also implement a relatively new procedure, which takes a permutation approach (see Moore and McCabe, 2006; Bowers *et al.* 2011; Johnson *et al.*, 2012). The aim of this approach is to use (most or) all of the available data and summarise the distribution for all possible scenarios (not just the best and worst). To do this, where the number of possible permutations is manageable, an overall mean effect size will be computed for each one. Where there are a very large number of possible permutations, a random sample will be selected using Monte Carlo simulation. This procedure will produce a distribution of standardised mean effect sizes and hence a more complete understanding of the likely overall impact of intervention.

Heterogeneity and sub group analysis

As noted, we expect the effect sizes to vary across studies. To quantify the degree of heterogeneity observed, we will calculate a Q statistic. This statistic is used to determine if any observed variation in effect sizes is likely to be above and beyond that which would be expected on the basis of sampling error alone (see Lipsey and Wilson, 2001). A statistically significant Q statistic therefore implies that there are systematic variations in effect sizes that cannot be explained by sampling error. Such differences would include, for example, variation in the contexts in which an intervention has been implemented.

If, as expected, significant heterogeneity is observed in the effect sizes, a moderator analysis will be conducted to see if variation in factors associated with the reviewed studies can explain this. As outlined above, some of the factors considered will be informed by our realist review of the available evidence, with the aim of ensuring that the analysis is theoretically informed.

As part of the moderator analysis, weighted mean effect sizes will be computed for each of the identified sub-groups. We will also calculate a Q statistic for each subgroup. In the event that the theoretically informed moderator variables explain the observed variation in effect sizes,

any remaining variation will be explained by sampling error alone, and the analysis will thus provide insight into (at least some of) the ingredients necessary for a successful intervention.

Publication bias

A well-documented issue that can compromise the reliability of the outcome of any meta-analysis is publication bias (e.g. Kicinski, 2014). Simply put, if evaluations that suggest positive outcomes of interventions are more likely to be published, there is a risk that any positive effect of intervention will be exaggerated. To determine the extent of a publication bias in our sample our studies, using a moderator analysis, we will first compare the mean effect size observed in published studies with that for unpublished studies. Next, we will produce a funnel plot, plotting the effect sizes against their standard error. In the event that there is no publication bias, the individual effect sizes should be more or less symmetric around the overall mean. If, however, there is an over-representation of studies that suggest an effect greater than the overall mean effect, this would suggest that publication bias is a distinct possibility. In this event, we will use the trim and fill method proposed by Duval and Tweedie (2000) to estimate the true effect size of intervention.

Outlier analysis

Outlying individual outcomes can distort the overall mean effect size estimate. Such outliers are particularly serious if they have extreme values and/or come from large studies that have been heavily weighted in the meta-analysis. The analysis will check for the existence, and where appropriate, the influence of such outliers. This will be done by visually assessing the presence of any extreme values from the forest plots. If it is apparent that potentially problematic outliers exist, mean effect sizes will be calculated both with and without the inclusion of the extreme value. This sensitivity analysis will establish whether mean effect sizes in such situations are robust and consistent or whether the outlier has a problematic influence over the results found.

Inter-rater reliability

This looks at the reliability of the data extraction process. Checks will be made to assess whether the subjectivity of the coding process has an influence over the information

(particularly that relating to outcome) that is extracted. For a random sample of studies, two coders will independently code the outcome measures and other findings. These will be compared and the level of agreement will be calculated. If at any stage of the process, or for any category inter-rater reliability is low, measures will be put in place to resolve any differences.

6. PLANS FOR UPDATING THE REVIEW

The authors expect to update the review every five years if funding exists.

7. ACKNOWLEDGEMENTS

We are grateful for comments on this protocol from Prof George Mair and Prof Martin Killias.

8. REFERENCES

Austin, J. and Hardyman, P. (1991) 'The use of Early Parole with EM to Control Prison Overcrowding: Evaluation of the Oklahoma Department of Corrections Pre Parole Supervised Release with EM', National Institute of Justice

Bales W. (2010), 'A quantitative and qualitative assessment of electronic monitoring report submitted to the Office of Justice Program', NCJRS, (Available at) http://www.ncjrs.gov/pdffiles1/nij/grants/230530.pdf

Baumer, T. L. and Mendelsohn, R. I. (1991) 'Comparing Methods of Monitoring Home Detention: The Results of a Field Experiment', American Society of Criminology.

Black, M. and Smith, R. (2003) *EM in the Criminal Justice System*, Australian Institute of Criminology No. 254.

Bowers, K. J., Johnson, S. D., Guerette, R., Summers, L. and Poynton, S. (2011)'Do geographically focussed police initiatives displace crime or diffuse benefits? A systematic review', *Journal of Experimental Criminology*, 7(4), 347-374.

Bulman P. (2013), 'Sex Offenders Monitored by GPS Found to Commit Fewer Crimes', *NIJ Journal*, 271; 22-25

Crowe A., Sydney L., Bancroft P., and Lawrence B. (2002), *Offender Supervision with Electronic Technology*, Kentucky: American Probation and Parole Association.

Davidoff, F. (2009) 'Heterogeneity is not Always Noise: Lessons from Improvement', *JAMA*, 302(23), 2580-2586.

DeMichele, M. and Payne, B. (2009*) Offender Supervision with Electronic Technology: Community Corrections Resource, 2nd ed.,* Bureau of Justice Assistance, US Department of Justice

Dixon-Woods, M., Agarwal, S., Jones, D., Young, B. and Sutton, A. (2005) 'Synthesising qualitative and quantitative evidence: a review of possible methods', *Journal of Health Services Research and Policy*, 10, 45-53.

Dodgson, K., Goodwin, P., Howard, P, Clewellyn-Thomas, S., Mortimer E., Russell, N. and Weiner, M. (2001) *EM of Released Prisoners: an evaluation of the Home Detention Curfew Scheme*, Home Office Research Study 222.

Duval, S. and Tweedie, R. (2000) 'Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis', *Biometrics*, 56(2), 455-463.

Eck, J. E. and Madensen, T. D. (2009)' Using signatures of opportunity structures to examine mechanisms in crime prevention evaluations', In J. Knutsson and N. Tilley (eds.) *Evaluating Crime Reduction Initiatives*, Crime Prevention Studies, vol. 24, 59-48. Monsey, NY: Criminal Justice Press.

Eck, J. E. (2006) 'Preventing Crime at Places', In *Evidence-Based Crime Prevention*, rev. ed., edited by L. W. Sherman, D. P. Farrington, B. C. Welsh, and D. L. MacKenzie. New York: Routledge.

Erez, E., Ibarra, P. R., Bales, W. D. and Gur, O. M. (2012) 'GPS Monitoring Technologies and Domestic Violence: An Evaluation Study', *NCJRS Report*, June 2012

Feinstein, A. (1998)' A problem of cogent subgroups: A Clinicostatistical tragedy', *Journal of Clinical Epidemiology*, 51(4), 297-299.

Finn, M. A. and Muirhead-Steves, S. (2002) 'The Effectiveness of EM with violent Male Parolees', *Justice Quarterly*, 19:2

Garland, D. (2002) *The Culture of Control. Crime and Social Order in Contemporary Society.* Oxford University Press.

Gibbs, A. and King, D. (2003) *Home Detention with EM: The New Zealand Experience.* Criminal Justice. Sage Publications. Vol. 3 (2): 199 – 211.

Glaser, D. and Watts, R. (1992) 'EM of Drug Offenders on Probation', *Judicature* 76(3):112-17.

Gough, D., Thomas, J. and Oliver, S. (2012)' Clarifying differences between review designs and methods', *Systematic Reviews*, 1-28.

Guerette, R. T. (2009)' The pull, push and expansion of situational crime prevention evaluation: An Appraisal of thirty-seven years of research', in N. Tilley and J. Knutsson (eds.), Evaluating Crime Reduction Initiatives, Crime Prevention Studies, Vol. 24.

Haverkamp, R., Mayer, M. and Levy, R. (2004) 'EM in Europe*, European Journal of Crime, Criminal Law and Criminal Justice*, Vol. 12/1, 36 – 45.

House of Commons Committee of Public Accounts (2006) *The EM of Adult Offenders: 62nd Report of Session 2005-2006,* House of Commons, The Stationary Office Limited: London

Horowitz, R. (1987) 'Complexity and contradiction in clinical trial research', *The American Journal of Medicine*, 82, 498-510.

Horowitz, R. (1995) 'Large-scale randomized evidence: Large, simple trials and overviews of trials: Discussion: A Clinician's perspective on meta-analyses', *Journal of Clinical Epidemiology,* 48(1), 41-44.

Johnson, S. D., Guerette, R. T. and Bowers, K. J. (2012) 'Crime displacement and diffusion of benefits', In B.C. Welsh and D.F. Farrington (eds.), *The Oxford Handbook of Crime Prevention*. Oxford: Oxford University Press.

Johnson, S.D., Birks., D, McLaughlin, L., Bowers, K.J. and Pease, K. (2007) *Prospective mapping in operational context*. Home Office Online Report 19/07. Home Office: London.

Jones, M. and Ross, D. L. (1997) 'Electronic House Arrest and Boot Camp in North Carolina: Comparing Recidivism', *Criminal Justice Policy Review*, Vol. 8. No. 4/97, pp 383 – 403.

Kicinski M (2014) 'Correction: Publication bias in recent meta-analyses', *PLoS ONE* 9(1): 10.1371/annotation/a65c0f61-eb99-42f0-828b-5a8662bce4f7.

Laycock, G. and Tilley, N. (1995)' Implementing crime prevention programmes', In M. Tonry and D. Farrington (eds.) *Building a Safer Society*. Crime and Justice: A Review of Research. Chicago: University of Chicago Press

Lipner, S. L. (1993) *Performance Audit of the use of EM within the community Control Programme Administered by Department of Corrections*, State of Florida, Office of the Auditor General.

Lipsey, M. W. and Wilson, D. B. (2001) *Practical meta-analysis*. Thousand Oaks, CA: Sage Publications.

Mackenzie, D. L. (2006) *What works in Corrections Reducing the criminal activities of offenders and delinquents,* Cambridge University Press.

Mairs, G. (2005) *EM in England and Wales. Evidence based or not?,* Criminal Justice, Sage.

Mairs, G. and C. Nee (1990) *EM: The Trials and Their Results.* Home Office Research Study 120, London: HMSO.

Mairs, G. and E. Mortimer, (1996) *Curfew Orders with EM.* Home Office Research Study 163, London: Home Office.

Marklund, F. and Holmberg, S. (2009) *Effects of Early release from Prison using Electronic tagging in Sweden,* Springer Science and Business Media.

Moore, D. S. and McCabe, G.P. (2006) *Introduction to the Practice of Statistics* (5th Ed.), WH Freeman.

National Audit Office (2006) *The EM of Offenders*, available at
https://www.nao.org.uk/report/the-electronic-monitoring-of-adult-offenders/

Paterson, C. (2007) 'Commercial Crime Control and the EM of Offenders in England and Wales', *Social Justice*, Vol. 34, Nos. 3 – 4 (2007 – 2008).

Pawson, R. (2006) *Evidence-Based Policy. A Realist Perspective*, London: Sage.

Pawson, R. and Tilley, N. (1997) *Realistic Evaluation*, London: Sage.

Payne, B. K., DeMichele, M. and Okufo, N. (2009) 'Attitudes about EM: Minority and Majority Racial Group Differences', *Journal of Criminal Justice*, 37, 155 – 162.

Payne, K. B. and DeMichele, M. (2011) 'Sex offender policies: Considering unanticipated consequences of GPS sex offender monitoring', *Aggression and Violent Behavior*, 16: 177-187

Perry, A. E., Weisburd, D. and Hewitt, C. (2010)' Are criminologists describing randomized controlled trials in ways that allow us to assess them? Findings from a sample of crime and justice trials', *Journal of Experimental Criminology*, 6(3), 245–262.

Peto, R., Collins, R. and Gray, R. (1995) 'Large-scale randomized evidence: Large, simple trials and overviews of trials', *Journal of Clinical Epidemiology*, 48(1), 23-40.

Petrosino, A., Boruch, R. F., Soydan, H., Duggan, L. and Sanchez-Meca, J. (2001)' Meeting the challenges of evidence-based policy: The Campbell Collaboration', *Annals of the American Academy of Political and Social Sciences*, *578*, 14-34.

Renzema, M. and Mayo-Wilson, E. (2005) 'Can EM Reduce Crime for moderate to high risk offenders?*', Journal of Experimental Criminology,* 1: 215 – 237.

Rothwell, P. (2005a) 'External validity of randomised controlled trials: "To whom do the results of this trial apply?"', *The Lancet*, 365, 82-93.

Rothwell, P. (2005b) 'Subgroup analysis in randomised controlled trials: Importance, indications, and interpretation', *The Lancet*, 365, 176-186.

Sackett, D.L., Rosenberg, W.M.C., Gray, J.A.M., Haynes, R.B. and Richardson, W.S. (1996) 'Evidence based medicine: what it is and what it isn't', *BMJ*, 312: 71–2.

Shapps, G. (2006) *The Tagging Game*, available online at http://www.shapps.com/reports/the-tagging-game.Pdf

Shute, S. (2007) *Satellite Tracking of Offenders. A study of the Pilots in England and Wales,* Ministry of Justice Research Summary 4.

Sidebottom, A. and Tilley, N. (2012) 'Further improving reporting in crime and justice: an addendum to Perry, Weisburd and Hewitt (2010)', *Journal of Experimental Criminology*, 8(1), 49-69.

Sidebottom, A., Tompson, L., Thornton, A., Bullock, K., Tilley, N., Bowers, K. and Johnson, S. D. (2015) 'Gating Alleys to Reduce Crime: A Meta-Analysis and Realist Synthesis' (under review)

Smith, D. (2001) *EM of Offenders. The Scottish Experience,* Criminal Justice, Sage Publications Vol. 1 (2): 201 – 214.

Stacey, T. (2006) 'Electronic Tagging of Offenders: A Global View', *International Review of Law, Computers and Technology*. 20(1-2), 117-121.
Sugg, D., Moore, L. and Howard P. (1991) *EM and Offending Behaviour – Reconviction Results for the 2nd year of trial of curfew orders,* Home Office Research Study 141.

Tilley, N. (2006)' Knowing and doing: Guidance and good practice in crime prevention', In J. Knutsson and R.V. Clarke (eds), *Putting Theory to Work: Implementing Situational Prevention and Problem-Oriented Policing*. Crime Prevention Studies, vol. 20, Monsey, NY: Criminal Justice Press.

van der Knaap, L. M., Leeuw, F.L., Bogaerts, S. and Nijssen, L.T.J. (2008)' Combining Campbell Standards and the Realist Evaluation Approach: The Best of Two Worlds?', *American Journal of Evaluation*, 29(1), 48-57.

Whitehead S., Kersley H. and Mohun J. (2013) *Justice for Sale- the privatisation of offender management services,* New Economics Foundation (Accessed 08 December 2015)
https://www.tuc.org.uk/sites/default/files/Criminal%20Justice%20Report.pdf