Canterbury Research and Theses Environment

Canterbury Christ Church University's repository of research outputs

http://create.canterbury.ac.uk

Please cite this publication as follows:

Tong, S. and Heaton, R. (2016) Evidence-based policing: from effectiveness to cost-effectiveness. Policing: A Journal for Policy and Practice, 10 (1). pp. 60-70. ISSN 1752-4520.

Link to official URL (if available):

http://dx.doi.org/10.1093/police/pav030

Contact: create.library@canterbury.ac.uk

Evidence-Based Policing: From Effectiveness to Cost-Effectiveness

By Robert Heaton and Stephen Tong

School of Law, Criminal Justice and Computing, Canterbury Christ Church University, North Holmes Road, Canterbury, Kent, UK. E-mail: robert.heaton@talktalk.net and steve.tong@canterbury.ac.uk

Abstract

Recent years have seen the development of quantitative studies into policing effectiveness, in particular, the 'evidence based policing' movement which has encouraged the use of randomized control trials in the UK and the USA. Despite their significance, such studies remain narrowly based in terms of their take-up by academic institutions and police forces.

This article charts the rise of evidence-based policing and considers whether it could be taken a step further, by developing consideration of police effectiveness into that of cost-effectiveness. The use of 'Quality Adjusted Life Years' (QALY) methodology in the UK in the arena of drugs approval for use by the National Health Service, is considered as a model which might be transferable to policing. It is concluded that there are substantial similarities. Providing that the improvements sought are realistic, there is real potential for the cost-effectiveness of policing methods to be assessed.

Introduction

Recently universities have experienced increasing engagement with policing through conducting research, generating publications, and hosting conferences centred around the topic of making the police more effective in carrying out their functions, of which crime reduction is perhaps the most obvious example.

Reiner's illustration of the developing research agenda (Reiner's five stages of police research include: consensus, controversy, conflict, contradiction, and crime control) charts the emphasis of research in policing in five different stages from the 1960s to the present day (Reiner and Newburn, 2008). Reiner describes the current research stage as 'crime control' with characteristics including intelligence-driven approaches, community-orientated strategies, and the use of analysts aimed at deterrence and more directed operational strategies. It is in this 'crime control' stage that 'crime science' and 'evidence-based policing' are firmly rooted. The last five or so years have seen significant, though as yet patchy developments in the police research enterprise in the UK and the USA. University-led research carried out by academics and police practitioners has seen new and significant, if narrowly channelled enthusiasm for quantitative studies, including the use of randomized control trials (RCTs). These are 'gold standard' research methods and therefore the potential for adoption of the results as practice policy is substantial.

From early studies, articulating the 'discovery of discretion' through to the development of problem-oriented policing (POP) and intelligence-led policing (ILP), debates have taken place amongst professional academics and academically minded police officers and staff, at the end of which similar conclusions have been repeatedly drawn in relation to experimentation in policing (Reiner and Newburn, 2008; Waddington, 1999). The police, moulded by the 'Thatcher Revolution' and 'New Public Management' (Golding and Savage, 2008, p. 736), are considered to be operationally overly goal-focused, so that sufficient rigour in methodology and evaluation are consistently abandoned in favour of the necessity to demonstrate 'success' (Cockcroft and Beattie, 2006). Hirschfield et al. (2014, p. 308) argue:

> The police subculture and organisational ethos have rarely encouraged positive learning from mistakes and failures and such salutary lessons are rarely sought out, disseminated or acted on.

Barriers to applying research to practice are not one-sided with academics in their turn agreeing that their use of abstruse terminology may have been unhelpful to the joint endeavour and publications presented in ways that are not particularly user-friendly to busy operational officers (Dawson and Williams, 2009). At some point in these debates, it may be suggested that the linking of research with practice in the medical world might be a helpful model, particularly in relation to the evaluation of drugs for use by the UK's National Health Service.

Further consideration of this proposition gives rise to some caution. The approval of a new medicine is not a binary decision based upon absolute effectiveness or non-effectiveness. It is the result of an assessment of effectiveness in function and over time, in relation to the cost of purchase per patient. Similar considerations might also apply to policing. An intervention may be effective, but only marginally so, or over a very short timescale, or it may be prohibitively expensive. Ideally, the effectiveness of a policing innovation would be measured, and an assessment made of whether its costs would justify implementation against a standard; for example, reduction in crimes per unit cost.

The remainder of this article will discuss these ideas in further detail, in the context of research experiments in crime reduction with reference to harm, cost, and effectiveness. The development of quantitative aspects of policing research, by police and by academics, is summarized, including recent moves towards 'evidence-based' policing. This is followed by a description of the assessment of new drugs by the National Institute of Care Excellence (NICE), in particular the use of the Quality Adjusted Life Years (QALY) system as a means to assess cost-effectiveness. Finally, the potential transferability of similar methodology to policing is discussed.

Quantitative research by academics

Historically, the tradition of police-related research by professional academics has been primarily qualitative. For example, the encyclopaedic Handbook of Policing contains 30 chapters describing police history, culture, organization, working methods, and themes of debate (Newburn, 2008). There is minimal quantitative information throughout. A comparison of 'intelligence led', 'community', and 'problem oriented' policing identifies that each has the expected benefit of 'reduced crime' (Tilley, 2008, p. 388), but the extent of this and whether or not the necessary actions might be cost-effective, is unclear.

Elsewhere, there are examples of innovations in crime reduction, Newark (USA) Safer Cities Initiative (1998–2005) and the Minneapolis hot spot patrols (1988–89), where 'before and after' results appear to be significant (Sherman and Weisburd, 1995; Kelling, 2005). The impact of these innovations if translated into wider policy is rarely discussed, for example, in terms of crime reduction across an entire police force area.

Studies have also sought to establish post-facto, the nature of linkages between crime reduction, police action, and environmental factors. For example, the reasons underlying the national crime reduction in the USA in the 1990s have been examined at length, though with few firm conclusions (Blumstein, 2006). In the UK, a Home Office study showed that numbers of thefts and burglaries were related to consumer expenditure (crime opportunities) and the number of young males available to commit such crimes (Field, 1999).

However, the use of rigorous quantitative data, to underpin policing initiatives from their outsets, has been rare. Major theories such as ILP and POP have been propounded with scant consideration of available quantitative information, and virtually no assessment of cost-effectiveness. The ILP movement gathered strength in the mid-1990s following the publication of the Audit Commission report 'Helping With Enquiries', in response to burgeoning levels of volume crime such as burglary and thefts to and from motor vehicles (Audit Commission, 1993). The report was centred around the notion that the curtailment of the activities of prolific offenders would have a disproportionately high impact upon crime levels, relying upon an earlier Home Office report that 7% of males were responsible for 65% of crime (Home Office, 1989).

Only brief consideration of this data is required, to realize that the 'seven per cent' refers not to the offending population, but to the male population at large. A medium-sized town of 100,000 population, half of which were male, would require 3,500 people to be targeted in some way, to gain the desired leverage over 65% of crime. Setting aside ethical considerations, this number is so far beyond the capacity of policing as to be meaningless, except to point to the opposite conclusion to that initially indicated by the data, which is that the direct influence of policing upon crime rates is extremely limited.

There is very little qualitative information in relation to the impact of POP, but again the strong indications are that it is very limited. For example, the UK's Cleveland Police claimed to be a 'national leader' in the field, after claiming to use the technique just 82 times in the course of the financial year 1999–2000 (Cleveland Police, 2000, p. 3) In the same year, the force recorded 65,185 crimes (Home Office, 2001 and answered 99,665 '999' calls (HMIC, 2001, p. 12). If each solved problem prevented ten crimes or ten '999' calls per year on the 82 occasions POP was used, its impact would be a reduction in the region of 1% of either crimes or '999' calls.

Similar conclusions were reached by researchers studying the impact of POP implementation in London, where researchers noted that:

> The most fundamental difficulty was the sheer number of people policed by a single sector and its officers . . . officers can hardly be expected to 'own' the problems of such a large number of people in any meaningful way (Dixon and Stanko, 1995, p. 178).

Similarly, Cordner's examination of the classic Newport News POP experiment revealed that:

> Had this tactical effort been judged by the total amount of crime rather than just thefts from autos, however, or if it had been judged on the basis of citywide data rather than target-area data, its impact might have been missed (Cordner, 1996, p. 198).

This key observation points to the conclusion that the exercise of theories such as ILP and POP may be very useful at a local neighbourhood level, particularly in respect of problems which exhibit a substantial degree of patterning caused by repeat offending or repeat locations. The limits of policing capacity make them a less-realistic option for the reduction of crimes or emergency calls on a larger scale, such as the whole of a police force area or a town within it. Force-level crime statistics reported to the Home Office every 3 months, are a less than useful proxy measure of police action.

Quantitative research by police officers

The last 20 or so years have seen the burgeoning of academic research by police officers, almost invariably associated with the obtaining of an academic qualification. This is offered by many universities and typically comprises a 2-year, part-time Masters-level degree. The research element is likely to cover a maximum timescale of 6 months and see the production of a dissertation

consisting of about 15,000 words. Although an intensive experience for the student, the process is relatively brief and terms of reference for research are necessarily tightly drawn.

The self-contained nature of such research means that there is little opportunity for further exploration of interesting results and following the award of the qualification, little incentive to publish. For example, research might investigate whether there is an association between the variables of police action, and reduced crime. Production of a statistical p-value of less than 0.05 would indicate a statistically significant association, and may become the main research result, although it would have little to say about the strength of the association and what the potential for crime reduction may be in practice. There is also scant evidence in the literature, of the baton of such research results being passed on for further investigation. The a priori suspicion must be that significant research results have never been drawn to the attention of policymakers.

Quantitative research in policing—recent developments

The last decade has seen limited but significant progress in quantitative research, owing to the convergence of several factors which saw the beginning of the 'evidence based policing' movement. In the UK, 2007 saw the appointment of Professor L.W. Sherman to Cambridge University's Institute of Criminology in 2012. The appointment saw increased emphasis upon quantitative research at the Institute, and in particular, promotion of the use of RCTs in an environment where these had been previously non-existent. The strength of RCTs is that they are able to negate unknown variables which may affect simple 'before and after' measurements.

In 2010, a group of UK police officers founded the Society for Evidence-Based Policing (SEBP), a body of police and academics devoted to the production and dissemination of evidence-based research. This was to be achieved by supporting research including access to police data, and by the provision of annual conferences.

In the same year, the former National Police Improvement Agency was developed into the new College of Policing which was formed with a specific government remit, of promoting 'what works' research. The College of Policing and SEBP set up databases of research projects, although it is unclear to what extent these are comprehensive or contain overlapping information.

In the USA, Professor David Weisburd, a long-term research partner of Sherman, founded the Centre for Evidence-Based Crime Policy at George Mason University, Washington, DC in 2008. The Centre produced another database of research studies, the Evidence-Based Policing Matrix. The Matrix was limited to those studies which reached fixed criteria of methodological rigour, namely RCTs, or quasi-experiments using matched comparison groups or multivariate controls (George Mason University, 2013).

The transatlantic alliance of these university departments attended by police leaders, backed by the SEBP and the College of Policing, has become a powerful force for the promotion of evidence-based experiments in policing (Neyroud and Weisburd, 2014). Beyond this core of activity, the majority of academic institutions have remained on the sidelines. There has been limited criticism of the 'evidence based' movement, to the overall effect that it requires 'some counterbalance and caution'. Greene argues:

> . . . there are many voices seeking to balance experimentation as applied to police interventions and drawing from a much wider array of theories and research methods that make the police contextual, that is, research that places the police in their environmental milieu replete with varying communications, interpretations, and meanings. (2013, p. 194)

Greene's point here is that there is undoubted value to experimentation but its value is enhanced when balanced with other approaches. Other criticisms point to EBP approaches as being slow and expensive, in an environment where police officers are often required to solve problems in a short time. Moreover, experience and skills are as important as formal scientific enquiry, in solving human problems. Furthermore, if action to solve a problem is effective, determining the precise causality may be of secondary importance (Sparrow, 2011).

NICE and the use of QALYs

The UK's National Institute for Health and Care Excellence (NICE) replaced the National Institute for Clinical Excellence in 2013, as a Non-Departmental Public Body. NICE is accountable to its sponsor, the Department of Health, but is operationally independent of government. The aim of NICE is to 'reduce regional variations in the availability and quality of health treatment, by issuing national guidelines' (NICE, 2014).

In relation to the provision of drugs, an independent NICE committee reviews evidence of effectiveness and further, whether or not the drugs are also cost-effective. This is achieved in relation to life-threatening illnesses, by measuring a drug's anticipated benefit to the patient in terms of QALYs. QALYs acknowledge that various drugs used to treat a particular condition may be different in terms of size of impact including side-effects (the quality factor), the timescale over which they are effective (the longevity factor), and their cost. The benefit is assessed against the cost of the drug in terms of '£per QALY'. In general, the drug should cost less than £20,000–£30,000 per QALY to be approved by NICE.

A drug may be assessed as producing a quality factor against a scale where 0 = dead and 1 = best possible health. For example, the patient may expect confinement to a wheelchair which might equate to a quality factor of 0.7. The drug may give the patient a longevity factor of 2 years of additional life at a total cost of £20,000. The QALY would be 0.7 x 2 = 1.4.

Cost-effectiveness would be £20,000 divided by 1.4 = £14,286 per QALY.

The evidence frequently includes the use of quantitative research techniques including NICE's preferred option, RCTs. However, this is not always possible such as when sample sizes are small or when interventions are complex and multifaceted (NICE, 2012). The findings are generally but not universally applied in the UK. For example, in 2010 the government initiated a Cancer Drugs Fund, to pay for drugs which had been rejected or not yet approved by NICE.

The use of QALYs is widely, but not universally accepted. Perhaps, most seriously, the quality factors relate to an 'average' patient, and have been determined by referral to disinterested parties such as students. In practice, the subjective experience of directly-affected patients may be very different. Furthermore, QALYs may discriminate against elderly patients, who in general have fewer life-years remaining than their younger counterparts. Moreover, the use of QALYs does not take into account, the impact on lives other than those of patients, such as carers.

QALYs are also used in North America, but to a lesser extent than in the UK. In particular, the Patient Protection and Affordable Care Act of 2010 ('Obamacare') prohibited the use of QALYs to establish a cost threshold for health interventions, following public concern about the prospect of 'death panels' (Longman, 2013).

The transfer of QALY principles to policing

In principle, assessment of the cost-effectiveness of crime reduction techniques is fairly straightforward. Each individual category of crime would be allocated a 'harm factor'. An 'effectiveness factor' of each type of policing intervention would be established by research evidence, including lifespan of that effectiveness. The benefit of the intervention would be calculated by multiplying the harm factor with the effectiveness factor. The cost of intervention would be divided by the resultant 'benefit factor', to assess the cost per unit of benefit.

Therefore, Cost benefit of intervention = Cost of intervention/harm factor x effectiveness factor

Intervention would only take place below a prescribed cost-benefit threshold.

Example:

Action will be approved below a threshold of £500 per benefit factor.

A serious assault is allocated a harm factor of 0.7. Specified action to prevent it is calculated to cost £100 and has an effectiveness factor of 0.5.

Cost benefit would be 100/0.7 x 0.5 = £286 per benefit factor. The action would be approved.

Graffiti is allocated a harm factor of 0.2. A proposed intervention has an effectiveness factor of 0.1 and also costs £100.

Cost benefit would be 100/0.2x 0.1 = £5000 per benefit factor. The action would not be approved.

The harm factor

The concept and use of 'harm' in policing have attracted contemporary debate. The term has been seen as attractive, because it offers the opportunity to give relative weight to the impacts of various crime types upon victims and the wider community. The concomitant problem is that its potential breadth of scope invites inconsistencies in interpretation and measurement. Beyond crime reduction, the role of police becomes increasingly difficult to define and the assessment of harm reduction becomes more problematic, particularly in respect of hidden societal problems (Ratcliff, 2015).

This approach focuses primarily on crimes the police know about and the harm inflicted by citizens and not the state. Harm can be inflicted by the state in terms of an unwillingness or ineffectiveness in dealing with specific crimes and therefore escalating the harm inflicted and this is explicitly related to police (in) effectiveness. Examples of historic major crimes or events revisited have included Stephen Lawrence's murder, the Hillsborough tragedy, the Rotherham child abuse scandal, and Jimmy Saville's serial abuse. Miscarriages of justice occurring because of police corruption, poor investigation, or judicial practices. More systematically, the police receive continued criticism around reporting rates and responses to hate crime, sexual offences, and domestic abuse not inspiring confidence for victims to report crimes against them. In an era where policy makers claim to be more victim centred and while engaging with restorative approaches to practice, it is not only the monetary value that can be controversial but variation in police responses to particular types of harm that becomes important to any consideration of effectiveness and the extent of harm inflicted.

If its scope is limited to crime reduction, the construction and application of a 'harm index' becomes more manageable. Attempts have included the monetization of various offences, and the use of opinion surveys of criminal justice professionals and others (Ratcliff, 2015).

Linkage of harm to sentences upon conviction of offences has emerged as the preferred basis of an index. This has the advantages that it is unnecessary to assign an artificial financial cost to non-property crime, and is broadly reflective of public opinion over time. There remains some debate over the precise nature of the linkage, in particular, whether the analysis of pre-sentence guidelines for courts should be preferred to those sentences which have been actually handed down.

A 15-point gravity index has been compiled by the Pennsylvania Commission on Sentencing as part of its guidance to the judiciary. This allows for some gradation of offences dependent upon the attendant circumstances. For example, the score for a burglary would be 5, 7, or 9, depending upon whether it took place in a home, and whether or not anyone was at home. Ratcliff used this index to compile District-level harm scores in Philadelphia in the years 2004–13 inclusive, comprising different crime types, traffic accidents, and stop-search activity which could damage police-community relations. The graphs demonstrated that while overall harm declined, there were variations from one district to another which had the potential to trigger different police action from that suggested by simple crime counts He concluded by hoping 'to see an expansion of harm-focused, intelligence-led, evidence-based policing' (Ratcliff, 2015, p. 180).

Sherman (2013) proposed a harm index which would be based upon the median number of days in prison for an offender first convicted of an offence, this measure being unaffected by previous convictions. The index would be compiled in a multistep approach, as follows:

1.Count the number of crimes of each type (A) in a given area.

2.Multiply (A) by the median number of day's imprisonment upon first conviction (B).

3.The product of A and B would be called the HST (harm subtotal) for each offence type.

4.Repeat steps 1–3 for each crime type.

5.Sum up all HSTs to produce the total crime harm (TCH).

6.Divide the TCH by the population of the area, to produce a standardized crime harm index (CHI).

A cautionary note was struck by Ratcliff by pointing out the degree to which weightings are subjective. For example, a homicide is weighted as twice the gravity of a robbery in the Pennsylvania system, but 128 times greater in a monetary value-based system (Sentencing Guidelines Council, 2006).

It is not clear how weightings for homicide and robbery would be compared in Sherman's proposed system, or how it would be applied in practice to produce harm factors for the offences. The UK's Sentencing Guidelines Council divides robbery into three categories of seriousness, depending upon the nature, degree, and results of force used. The range of sentences vary markedly, that is, 0–3 years custody for the least serious category compared with 7–12 years for the most serious (Sentencing Guidelines Council, 2006). Lesser sentences apply to young offenders. In respect of murder, Schedule 21 of the Criminal Justice Act 2003 divides the offence into four degrees of seriousness for adults, with an additional category for young offenders. Minimum sentences of imprisonment for adults range from 15 years to whole-life terms. Moreover, police-recorded categories of offence do not conform to those used for sentencing purposes. For example, robbery is divided only into business property and personal property. All in all, the compilation of a harm index on the basis of sentencing guidelines may not be straightforward.

The effectiveness factor

Quantitative studies in policing frequently seek to demonstrate whether there is a relationship between cause and intended effect, such as whether specified police intervention reduces crime levels. The results usually take the form of a statistical table in which the p-value is used to test whether the research's hypothesis should be accepted or not. For example, a p-value of 0.05 would denote that there is a 5% probability that the effect of intervention was nil and the study's observations have arisen by chance. Therefore, a large-scale study could produce a small p-value although the effectiveness of police action may be marginal. A low p-value implies that there is a relationship between variables, because there is a low probability that the results were produced by chance. It has nothing to say about the strength of that relationship, for which 'effect size' statistics are required. Ideally, quantitative studies would produce complementary statistics in respect of both statistical significance and effect size. This is rare in policing-related research literature. Studies which conclude with statistical tables may demonstrate that police action is effective, but are almost invariably silent on the subject of cost-effectiveness.

This additional step taken is not necessarily difficult, as demonstrated by the report of 'Operation Beck', the first randomized control study of policing hotspots in England and Wales. This was initiated in conjunction with British Transport Police in 2011, with the intention of reducing incidents requiring police attention and the number of reported crimes, on London Underground platforms. Following analysis of the problem, the experiment applied 15-minute patrols, 4 times per shift on 4 days of the week, to 57 hotspots. No action was applied to a further 58 'control hotspots' which had experienced very similar crime levels.

The primary results were that calls for service in the hotspots were reduced by 21%, and crime was reduced by 15%, compared with the control hotspots. A table of descriptive statistics produced significant p-values. Significantly, the study went further by also producing a cost-benefit analysis. This showed that the gross cost of the operation was £6,000 per prevented crime and £200 per prevented call for service. However, the marginal additional cost of items such as police overtime was considerably lower, at £250 per prevented crime and £8 per prevented call for service (Ariel and Sherman, 2012).

When a drug is evaluated for possible use by the NHS by means of QALY calculations, the cost of the drug is used, without including the wider costs of administering it or indeed those of running the entire NHS. To take into account such costs, would introduce further arguable variables and dampen the difference between the costs of different treatments. In common with the police service, an organizational commitment has already been made to meet overhead, including day-to-day staffing costs. Therefore, it is arguably reasonable to judge effectiveness in terms of marginal additional costs.

The conversion of the results into an 'effectiveness factor' for the purpose of cost-benefit calculations is more difficult. Crime reduction of 15% could be used as an effectiveness factor of 0.15 as a rudimentary solution, but this takes no account of the crime reduction's longevity. NICE calculations of QALYs measure increases in life expectancy, which have well-defined end points upon the death of patients. This is not the case with crime reduction. Ideally, follow-up data would be required and some means found to incorporate this into the calculation. As indicated earlier in this article, the time-limited nature of much police research weighs against this possibility. Moreover,

the effect of 'diffused benefits' of crime reduction outside the target area may be difficult to quantify.

Discussion and conclusions

The last half decade has seen significant developments in quantitative policing research in the UK and the USA, through improvements of methodological rigour and in particular, the introduction of RCTs. These developments have been initiated by a small number of academics, and taken up with enthusiasm by practitioners with an interest in improving professional practice. If this somewhat limited base expands to the point of being routine university department activity, there is potential for the assessment of police effectiveness to be revolutionized. A reasonable next step, particularly at a time of public sector austerity, would be for assessments of effectiveness to be developed into those of cost-effectiveness. This article has shown that evaluation of cost-effectiveness has inherent methodological problems, particularly if a system were sought which required a comparison to be made of different offences, enabling one type of crime problem to be preferred for action over another.

A harm index would be necessary, to take account of the differential impacts of crime types. Its construction would require the making of subjective judgements which would undoubtedly be inaccurate in relation to some victims and communities. There is some consensus that a harm index should be linked to sentencing, although attempts to date have produced significantly different weightings.

A measure of effectiveness of police action would also be required and at its simplest, this might be the extent of crime reduction, for example, 30% reduction would equate to an effectiveness factor of 0.3. This neglects the longevity of impact and longer term work would be needed to establish this and also the effects of benefits of diffusion outside the area subject of research.

Further objections can be raised to the notion of measuring cost-effectiveness in policing experiments. Policing is predominantly a reactive activity, whether it is the need to respond to calls for assistance, or to attend high-volume crime, or to deal with serious incidents such as murder to exacting standards. The experience of intelligence-led and problem-oriented policing is that they can be locally effective, but are too resource-intensive to 'get on top' of high-volume problems on a wider scale. The impact of austerity on police resources is likely to decrease the proportion devoted to proactivity and therefore Ratcliff's 'harm-focused, intelligence-led, evidence-based policing' may amount to little more than a harm-focused, intelligence-led, evidence-based rump of the whole.

Examination of the QALY system provides not only a counterweight to the difficulties of measuring the cost-effectiveness of policing, but also a reality check. Returning to the primary aim of NICE, QALYs support its intention to 'reduce regional variations in the availability and quality of health treatment, by issuing national guidelines'. In the hard-science medical world, this ambition falls far short of achieving academic exactitude. Improvement is sought rather than perfection. While problems with QALYs are acknowledged, there is neither a consensus around an alternative, nor any serious suggestion that QALYs should be replaced by the random outcome of a postcode lottery.

Improvements are sought in the policing and medical worlds, against a background of heavy reactive pressures. In both spheres, local judgements and political pressures may from time to time, cause research-produced priorities to be overturned. The similarities go further. The majority of objections to QALYs mirror to a remarkable degree, those of measuring police cost-effectiveness. Assessments

of relative quality of health conditions are subjective, as are those of harm suffered by crime victims. The construction of health and harm indices relate to an 'average' person which may not correspond to the experiences of individuals. The construction of such indices according to different criteria may produce wide variations in results. The difficulty in measuring diffused crime reduction benefits is mirrored in the failure of QALYs to take into account patients' carers.

The use of QALYs in medicine demonstrates that providing that improvement rather than perfection is sought and reasonable consensus can be reached where judgement is required, all of these flaws can be overcome or accommodated. The evidence-based policing movement has initiated the process of assessing the effectiveness of policing innovations, which hitherto has been conspicuous by its absence. The cost of policing as a public service makes desirable, further development of research into the arena of cost-effectiveness.

The application of QALYs by NICE also highlights the stark contrast between the application of research in the medical and policing worlds. In the former case, a stream of research results is forced into the NICE filter, under the constant commercial pressure of the necessity to recoup the costs of research. Policing has no analogous process and research results exist in something more akin to an autonomous, free-floating state. It seems clear that a necessary further step towards cost-effectiveness in policing is the central coordination of such research. This is currently limited on both sides of the Atlantic to a few police forces working with an even smaller number of universities. The reason for limited take-up is unclear, but may be a consequence of preoccupation with immediate pressures, combined with maintenance of the status quo and the absence of incentive for change. Meanwhile, central criminal justice agencies appear to be circulating the results of research efforts, rather than exercising leadership by influencing their direction, clarifying ambiguities, promoting those which are most cost-effective and assisting their implementation. Until such mechanisms are put into place, the implementation of evidence-based policing research, whether at the level of effectiveness or cost-effectiveness, will be limited.

References

ACPO (2013). *ACPO Youth Offender Case Disposal Gravity Factor Matrix*. London: Association of Chief Police Officers.

Ariel B., and Sherman L. (2012) *Operation "BECK" Results from the First Randomised Controlled Trial on Hotspot Policing in England and Wales*, 5th International Evidence Based Policing Conference 9 - 11 July 2012, available from http://www.crim.cam.ac.uk/events/conferences/ebp/2012/beckrctresults.pdf, (accessed 28th May 2015).

Audit Commission (1993). *Helping with Enquiries*. London, Audit Commission.

Beattie I., and Cockcroft T. (2006). Square Pegs and Round Holes: Performance Measurement in the Police and Prison Services, *Prison Service Journal* 168: 39–44.

Blumstein A. (ed) (2006). The Crime Drop in America. Cambridge: Cambridge University Press.

Bryant R., Roach J., and Williams E. (2015). *Crime and Intelligence Analysis through Partnership* (CIAP), Final Report prepared for the College of Policing, Canterbury Christ Church University and University of Huddersfield.

Bryne S., and Pease K. (2008). 'Crime Reduction and Community Safety.' in T. Newburn (ed) *Handbook of Policing*, 2nd Edition, Cullompton: Willan Publishing.

Cleveland Police (2000). *Chief Constable's Annual Report 1999–2000*, Middlesbrough: Cleveland Police.

Cope N. (2004). Intelligence Led Policing or Policing Led Intelligence: Integrating Volume Crime Analysis into Policing, *British Journal of Criminology*, 44: 188–203.

Cordner G. W. (1996). 'Evaluating Tactical Patrol, in Hoover L. T.' *Quantifying Quality in Policing*, Washington D.C.: Police Executive Research Forum, pp. 207–242.

Dawson P., and Williams E. (2009). Reflections from a Police Research Unit – An Inside Job Policing: *A Journal of Policy and Practice* 3(4): 373–30.

Dixon B., and Stanko B. (1995). Sector Policing and Public Accountability, *Policing and Society* 5: 171–83.

Field S. (1999). *Trends in Crime Revisited*, Home Office Research Study 195. London: Home Office.

George Mason University (2013). *Inclusion Criteria & Methods Key*. http://cebcp.org/evidence-based-policing/the-matrix/inclusion-criteria-methods-key/ (accessed 18 July 2015)

Goldstein H. (1979). Improving Policing: A Problem Oriented Approach, *Crime and Delinquency* 25: 236–58.

Hirschfield A., Ekblom P., Armitage R., and Roach J. (2014). Holding the Line: The Sustainability of Police Involvement in Crime Prevention. In Brown J. M. (ed.) *The Future of Policing*. London: Routledge, pp. 299–316.

HMIC (2001). *Inspection Report: Cleveland Police 2000–2001*. London: Stationery Office.

Home Office (1989). Criminal and Custodial Careers of Those Born 1953, 1958 and 1963, *Statistical Bulletin 32/89*, London: Home Office.

Home Office (2001). *Criminal Statistics in England and Wales 2000*. London: Home Office.

Kelling G. L. (2005). Community Crime Reduction: Activating Formal and Informal Control. In Tilley N. (ed.), *Handbook of Crime Prevention and Community Safety*. Cullompton: Willan Publishing, pp. 107–42.

Laycock G. (2005). Defining Crime Science. In Smith M. J., Tilley N. (eds), *Crime Science: New Approaches to Preventing and Detecting Crime.* Cullompton: Willan Publishing, pp. 3–24.

Longman P. (2013). *The Republican Case for Waste in Health Care*. http://thehealthcareblog.com/blog/2013/03/08/the-republican-case-for-waste-in-health-care/#more-59128 , (accessed 20 July 2015).

Neyroud P., and Weisburd D. (2014). Transforming the Police through Science: Some New Thoughts on the Controversy and Challenge of Translation. In Weisburd D., Lum C. (eds), *Translational Criminology*. Washington DC: George Mason University.

NICE (2012). *Methods for the Development of NICE Public Health Guidance*, 3rd edn. London: National Institute of Health and Care Excellence.

NICE (2014). *Who are we*. https://www.nice.org.uk/about/who-we-are (accessed 29 May 2015).

Ratcliff J. H. (2015). Towards an Index for Harm-Focused Policing, *Policing* 9 (2): 164–82.

Sentencing Guidelines Council (2006). *Robbery – Definitive Guideline*. London: Sentencing Guidelines Council.

Sherman L. W. (2013). Targeting, Testing and Tracking Police Services: The Rise of Evidence-Based Policing, 1975–2025. In Tonry M. (ed.), *Crime and Justice in America, 1975–2025, Vol. Crime and Justice 43*. Chicago: University of Chicago Press.

Sherman L., and Weisburd D. (1995). General Deterrent Effects of Police Patrol in Crime Hotspots: A Randomized Controlled Trial, *Justice Quarterly* 12: 625–48.

Sparrow M. K. (2011). *Governing Science: New Perspectives in Policing*. Washington, DC: Department of Justice, National Institute of Justice.

Tilley N. (2008). Community Policing, Problem-Oriented Policing and Intelligence-Led Policing. In Newburn T. (ed.), *Handbook of Policing*, 2nd edn. Cullompton: Willan Publishing, pp. 373–403.