

Health Development Agency
Public Health Evidence Steering Group
Methodology Reference Group

Starter Paper

Methodological Problems in Constructing the Evidence Base in Public Health

Mike Kelly, Catherine Swann, Amanda Killoran, Bhash Naidoo, Elaine Barnett-Paige, & Antony Morgan

9 August 2002

Summary:

As part of its core function, the Health Development Agency is required to map and synthesize the evidence base for the effectiveness of public health interventions to improve health and reduce health inequalities. This activity has a number of challenges associated with the assessment of evidence, concepts of public health, and modes of transfer of evidence into practice. This paper outlines some of these challenges, and presents the brief for the Methodologies Reference group to collaborate with the HDA in developing new and innovative means of overcoming these challenges.

Introduction

The Health Development Agency (HDA) was established as a Special Health Authority in 2000, following the publication of the *Saving Lives: Our Healthier Nation* White Paper (Department of Health, 1999). One of the Agency's core functions is to build the evidence base in public health and health improvement, particularly with regard to the reduction of health inequalities. The importance of this function has since been reinforced in a number of official publications (Department of Health, 2001).

To fulfill this function, the HDA has developed a web-based database called Evidence Base (EB). Its purpose is to disseminate the best available evidence on what works to improve health and reduce health inequalities. The database, which focuses on review-level data, sits on the web site HDA Evidence Base, (HDA EB) which includes other resources and links for evidence based public health practice. This database and site were launched early in 2001, and will grow to bring together all of the best evidence on public health improvement as part of the larger Public Health Electronic Library (PHEL) project. Work to link to this electronic platform also underway includes the development of knowledge about the effective implementation of interventions and systems, and health impact assessment. 'Healthpromis' is a parallel HDA database, a bibliographic resource to compliment and expand upon the information included in HDA EB.

A number of topics, including prevention of accidental injuries, drug and alcohol misuse prevention, reducing teenage pregnancy, controlling the spread of HIV/AIDs, smoking prevention, and obesity reduction have been selected for inclusion in the HDA EB programme for the current year, and further topic reviews are planned. The materials that will constitute the database include pdf files of recent systematic and other reviews of effectiveness (if they are available in the public domain), bibliographic information on relevant research, and topic-specific briefing documents - based on collation and synthesis of review level data - prepared by the HDA's Research & Information Directorate. These will be called *Evidence Based Briefing Documents*. The website may be viewed on <http://www.hda-online.org.uk/evidence>

This paper introduces current approaches to collating and synthesizing public health evidence of effectiveness for inclusion in HDA EB, and the methodological and conceptual challenges that the Methodologies Reference Group are being asked to consider. At this early stage, the approaches are developmental and it is anticipated that the methodology will be revised and refined as the Reference Group and staff generate new theory. Parallel projects are currently being undertaken to develop suitable search strategies for identifying the evidence, and effective modes of transfer of evidence for effectiveness into public health practice.

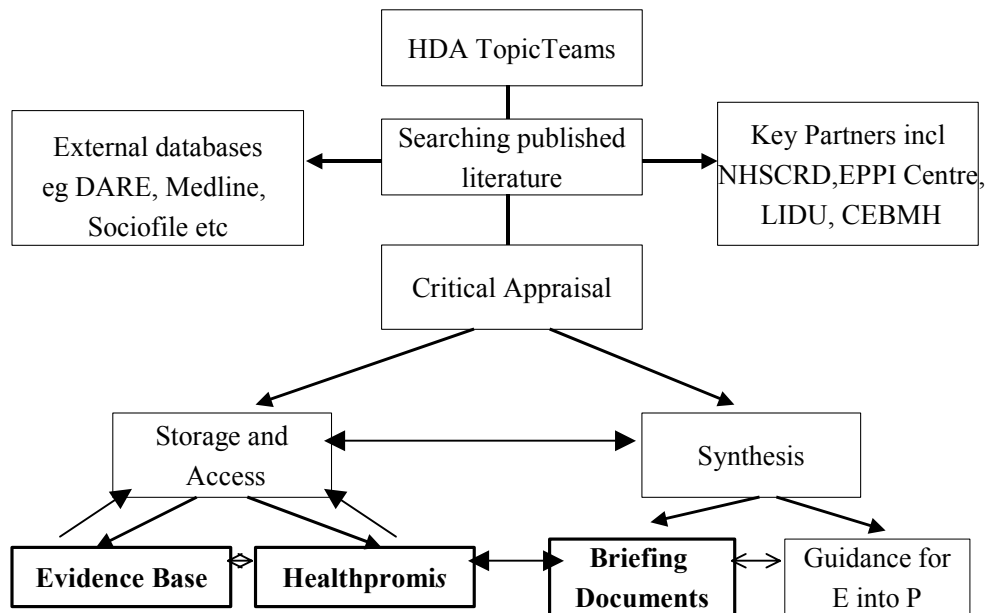
Collating the evidence: an overview of the process

Supporting HDA EB is a multi-faceted programme of work aiming to bring together all relevant evidence for the effectiveness of interventions to improve

health and reduce health inequalities, providing vital information for effective public health practice. The information contained on the database and website is aimed at public health policy and decision makers, practitioners, and researchers. The work plan within this programme is three-fold: The first stream of projects (including the creation of a Public Health Evidence Steering Group (PHESG) and its Methodologies Reference Group) is building the methodological and epistemological structures for the collation and interpretation of evidence for public health. The second stream of projects is engaged in mapping and synthesizing review-level evidence in priority topic areas. The third stream is concerned with electronic and web infrastructure of the site - developing and maintaining a web gateway to international sources of evidence, managing and updating the database itself, developing quality assurance protocols for the site's form and content, and linking in with the PHEL project. This paper, and the methodologies working group, will be concerned with streams one and two. The process is represented in figure 1. However, as this is the first year of the project the process is still in development and the method we present here is what we are currently working towards.

Figure1. Process of Building the Evidence base

How we will achieve our aim



Topics for inclusion are determined at the start of the business year, influenced by a number of factors including national public health targets and priorities, forthcoming and recent government strategies and frameworks, and sometimes

influenced by key partnerships. Once a topic has been selected, a project team is formed of research and public health specialist staff. Key national documents are identified, such as relevant national service frameworks and policy documents, to inform the guiding research questions that the team will pursue, to develop topic-relevant search terms, and to identify potential Reference Group members. The Reference Groups consist of research, public health and government topic specialists, and are convened to monitor the progress of the project team, inform the content of the brief, and review the final document.

A set of core search terms, a process for identifying topic-specific search terms, and a protocol identifying relevant databases to aid staff in implementing searches has been developed by HDA Research and Development and Information specialists, drawing on work carried out by the CRD at York, the Cochrane Collaboration, and Health Evidence Bulletins Wales (see Annex 1).

Once searches have been carried out, the team will scan the results (title and abstract) to identify all relevant review-level data that focuses on effectiveness. Two or more team members, who must agree that a paper is relevant, make selections. Selected papers are obtained, read and critically appraised again by two more more members of the team (a third reader will always be used where two readers cannot agree on the merit of a paper) using a critical appraisal (CA) tool developed specifically for this purpose (see Annex 2). Review level data is taken to mean systematic reviews of effectiveness, systematic reviews and reviews of effectiveness, meta analyses, literature, policy and narrative reviews, and all publications concerned with review and synthesis of research that is concerned with the research questions (see Annex 3 for more details).

The teams then re-read those reviews that have survived the CA process, and prepare an *Evidence Based Briefing Document* for each topic. The documents contain:

- a synthesis of the evidence
- discussion of the strengths and weaknesses of the evidence
- identification of any gaps in the evidence base
- an assessment of how up to date the evidence is
- identification of consensus or dispute around the evidence
- recommendations for new primary or secondary research, and broad policy implications

Where appropriate, each *Evidence Based Briefing* will inform the development of *Guidelines on Practice*, produced by the HDA. It is also anticipated that HDA will produce condensed, user-friendly summaries of the briefing papers, aimed specifically at practitioners.

The evidence pool

The process of collating the evidence and producing briefing papers on each topic is not without challenges. A description of the potential empirical (and practical) problems that need to be overcome in this project begins with a consideration of the very considerable empirical and practical problems that influence the narrow evidence base drawn from randomized controlled trials and other types of *quantitative* work. These are described in detail in work emanating from the Cochrane Centre in Oxford (Egger, *et al* 2001). The principal pitfalls arise from bias in the original research and bias introduced in the review process itself. The sources of bias have to be carefully assessed. Beyond the gold standard of systematic review of randomized controlled trials, reviews of cohort and observational studies are fraught with problems and meta analysis has its own particular difficulties (Egger, *et al* 2001). The systematic review of quantitative materials is a highly technically complex activity. However, once the unit of analysis shifts from the apparent *facts* presented by quantitative analyses to a broader range of data in all its potential forms, the practical and technical problems are even more considerable. Issues encountered in synthesizing non-experimental evidence, and in process reviews in areas such as change management, have their own difficulties (e.g. Mays *et al*, 2001). Measurement, validity, and reliability are probably even more contested in qualitative than in quantitative work, and bias more difficult to control.

In summary, the systematic and other reviews of effectiveness readily available to inform the development of the first *Evidence Based Briefing Documents* for HDA EB tend to be weighted towards a relatively narrow spectrum of potential evidence - that which is mostly drawn from randomized controlled trials and / or sits easily within traditional evidence hierarchies. Other types of methodological approaches - especially, but not exclusively, qualitative work - tend to be underrepresented in reviews of effectiveness. As the paucity of well-evaluated public health interventions has been observed elsewhere (Health Education Board for Scotland, 1996), this bias has the effect of further narrowing what is in many topic areas an already small pool of evidence. Yet even Cochrane (1971) himself argued that randomized controlled trials, for example, make a *contribution* to decisions about care and treatment - an observation that underlines the need for frameworks of quality and practice that can embrace the discipline of public health in its broadest sense.

The initial experience of collating the evidence base for the current year's priorities indicates that there is little qualitative research that directly assesses the impact or effectiveness of public health interventions and policies. This may be due to a number of factors: the databases and sources which have been searched may actively select against this type of data, this type of data may not answer the types of questions currently being asked, or the search terms do not facilitate its identification. However, the poor fit of qualitative work into current approaches to assessing effectiveness remains problematic because it may also exclude new and innovative methodologies (such as action research, where

impact and effectiveness may be central) from consideration. Because of these issues, there is a need to develop a process equivalent to the current methodology for locating, grading, synthesizing and interpreting data from qualitative, process, action and practice work.

A related problem is the gaps in the quantitative database. There are problems and topics which have either never been subject to the gaze of the researcher, or have only been studied in technically flawed ways, for example by poorly designed studies. This is particularly evident when the attempt is made to narrow the focus to a UK-setting, or on certain population sub-groups (which is often necessary, given the current emphasis on inequalities). In some of the areas where there is no work or only flawed work, qualitative data may fill some of the gaps. However, with respect to the flawed work, a very important lesson to emerge from the efforts of such groups as Cochrane is the necessity of rooting out poor research, or reviews of little value. From time to time this has caused considerable ill feeling towards the Cochrane approach (Mann 1990). It will be no less painful when the attention of reviewers from HDA extends to qualitative or other types of work and finds some of it wanting.

A final issue is the relatively 'poor' quality of the available quantitative evaluation data, which often fits only into the lower-end of traditional evidence hierarchies. Because of lack of fit between many community-based and public health interventions and rigorous experimental design, the field of available data, and thus the findings of relevant reviews and their applicability, may be limited.

Epistemologies of public health

Public health is by nature a multi-disciplinary concern. In collating the evidence base, researchers will draw upon work from sociological, psychological, anthropological, and medical traditions, to name but some. Within each of these areas exist various epistemological positions, many of which will be in direct philosophical conflict with others in the discipline of Public Health. There is a very clear epistemological problem attached to attempts to synthesize data collected using methods drawn from objectivist and subjectivist traditions. Most qualitative work starts from the assumption that the most important activity that the researcher is engaged in, is illuminating the intellectual categories which ordinary people (subjects) use to understand and function in the world in which they live. This approach tends to privilege the categories used by ordinary people. However, some proponents of this approach go further and deny the validity of the categories of science (which purportedly studies the material and social world as an object) as a means of understanding it. *In extremis* such arguments can descend into relativism and/or the rejection of the possibility of science at all. Not all researchers working in the subjectivist tradition (and there are many sub genres of such work) get anywhere close to nihilism or relativism. However, the tension between the view of reality as something which is socially constructed having no more permanence than the streams of consciousness of the subjects that created it, and a view which posits the existence of a real world having an

objective existence independent of the process of observing it, is an intellectual fault line which the Methodology Reference Group will have to traverse.

The tension is a very old philosophical problem. It can be found in the writings of Kant, Weber, Parsons, Habermas and many others (Callinicos, 1999). Sadly, although the philosophical arguments are subtle and complex, sometimes the arguments are reduced to unhelpful attacks, usually representing the position that is disagreed with, as logically impossible or plain stupid.

The potential evidence base in public health and health improvement draws upon a very wide variety of methodological traditions. The problems of inequalities in mortality and morbidity, and finding appropriate interventions to reduce inequalities, are too pressing a task to deny the possibility of objectivism or subjectivism, if they might help. The importance of finding ways of understanding the differences implied in the methodologies and of using this understanding to work with the evidence is primary - in collating the evidence base for public health, we are faced with the challenge of developing an epistemological position that allows us to acknowledge a variety of intellectual and practical approaches to the nature of truths and reality, and turn these into something that is useful and applicable for practitioners in the field.

There are a number of potential solutions. One is to posit that the objective material world does have a reality and does have properties and relationships in it, which exist independently of the scientist's (or lay person's) method of observing it. The scientist's methods of observing it may more or less provide an understanding of the mechanisms of that material world although it is an empirical question as to how accurate those understandings may be. This is not incompatible with the view that ordinary people make sense of and understand the material and social world in a variety of ways and use all sorts of intellectual categories to do so, some drawn from science. The understanding of those categories is a legitimate scientific activity too. However, just as the observer of the real material world is prone to error and misunderstanding, so too is the ordinary person and indeed the social scientist trying to understand that process. In other words both positions can exist logically without denying the legitimacy of the other, although different types and levels of analysis are involved.

It is important to acknowledge the different levels of analysis and to leave open the question, at least at this stage, of whether evidence drawn from the different traditions can be aggregated. The starting point is the recognition that data from the different traditions provide different perspectives. The concern is not to determine that one perspective is right and one is wrong, but to ascertain the degree to which the different perspectives can be validated and repeated, and are not idiosyncratic. This is not to argue that any perspective is therefore valid. It is to argue that the methodological task is to find means of evaluating research from whatever tradition it comes, according to agreed criteria of acceptability,

regardless of its theoretical or methodological origins. Is it valid, reliable, auditable and replicable?

The argument about different perspectives is not a rehash of the argument for triangulation, or not at least the version of triangulation that assumes that a truer answer or indeed a true answer to a problem will emerge because multiple methods are used. It should not be assumed that differences in perspective will produce similar or the same answers. It is just as likely that they will not. From a synthesis and policy point of view the practical question is what to do when in spite of valid, reliable, auditable and replicable work, different answers emerge from the different traditions.

Models of health production

There is a third set of potential problems. These relate to the way in which objects and phenomena within our understanding of public health are thought to relate to one another. The idea that health and ill health are determined by factors which are more than those which the health services do and deliver, has wide currency (Department of Health, 1999). Yet the mechanisms and pathways by which these wider determinants interact of health actually determine health has yet to be adequately articulated.

The notion of wider determinants of health is a helpful antidote to a narrowly biologicistic understanding of the aetiology of illness. However, it raises at least several further problems that are germane to the work of the Methodology Group. The wider determinants approach sometimes transmutes into a thinly disguised attack on the medical profession via an attack on the medical model. This is most unhelpful not least because it can lead to a position which is merely a generalized critique of social structures, without acknowledging the profoundly important benefits that biomedicine has provided to the human race in the treatment of many acute and chronic diseases. It also constructs a medical model that seems to have little real currency in medical practice. Of course medical power and professional self-serving interests are real enough. However a critique of the medical model and generalized criticisms of society are no antidote, and nor do they explain what the alternative social model might look like. A plausible alternative is required. That there are wider determinants of human health and disease and that the human body individually and collectively is susceptible to the damaging onslaught of geological, physical, social, political, economic and behavioural agents as well as bacterial and viral ones, is the basis of public health. How precisely these wider determinants work, the nature of their pathways of causation, and their interactions with the human psyche are not well understood, although many careful empirical studies offer a variety of clues. The testing methodological task is to be able to explore and map what is known and what is *not* known about these determinants, and indeed perhaps move away from the language of determinism altogether. This is to ensure that a facile attack on the medical model does not give way to an equally facile account of complex social reality. It is also to encourage the incorporation of the important

developments in social theory in recent years into public health thinking. The sophisticated accounts of social life developed by for example Giddens (1982) and Habermas (1987) and recent developments in theories of social stratification have very important roles to play (Graham, 2000).

A related issue here is how to judge the effectiveness of interventions at reducing health inequalities, within the pool of data that we obtain through our searches. Our early work indicates that very few studies explicitly address this in their methodology or analysis, and within those that do there is considerable variation in modes and units of analysis.

The politics of public health

There is a considerable body of opinion in the social scientific world that would reject the premise of the evidence based project upon which HDA EB is founded, and would be unsympathetic to the presentation of the issues in this paper. The line of attack is twofold. All science, including social science is a form of power and control, and the evidence base of public health and health improvement is nothing more than a sophisticated version of that means of domination, rendered all the more reprehensible because it is dressed in the language of reduction in inequalities (Lupton, 1995). Science, and perhaps in particular meta-analytic science that produces specific practices and activities, can be argued to actively create and discipline the populations it professes to help (Foucault, 1973; Turner, 1992).

The second criticism might be that HDA EB is some revival of positivism, which threatens to suppress the authentic voice of the victims of society, and therefore should be rejected in favour of “telling it how it really is”. There is actually no methodological or other argument, which can be used to very much effect to convince the proponents of such views. The Methodology Reference Group will need to acknowledge the potential barbs that will derive from these sources. However, the view within the R&I Directorate at HDA is that this ideological opposition is essentially disingenuous. It masquerades as a radical position politically, but in fact is fundamentally conservative. All knowledge may be power, or at least potentially powerful, but it does not follow that anyone and everyone doing science is a dishonest servant of the interests of global capitalism. *A priori* there is no reason why knowledge should not be used for the benefit of society just as surely as it can be used for its disbenefit. In the final analysis it is easier to retreat into a position, which decries any attempt to improve the health of the public as cynical manipulation, and do nothing about it. However, there are potential solutions - epistemological positions on the nature of ‘reality’ (in this case, the production of health) do exist that may allow us to circumvent these arguments - and the task of the Methodology Group is to map the territory in order to help in the task of reducing inequalities in premature mortality in men, women and children and reducing the burden of morbidity afflicting disproportionately the most disadvantaged in our society.

Conclusion

In summary, the success of our work will be determined by our ability to overcome, or work with, the following issues:

- current tools for assessing evidence are biased towards the upper-end of traditional evidence hierarchies
- limitations of the available data pool
- muddy theory around the production of public health
- tensions within the data pool in the way that health is conceptualized and researched
- lack of framework for / data judging the effectiveness of interventions to reduce health inequalities
- political factors within the field around issues of science and control

There is considerable national and international attention currently focused on the implications of broadening the evidence base to include additional types of data, such as qualitative research. Institutions and collaborations engaged in this area include the NHSCRD at the University of York, EPPI-Centre at the University of London, the Cochrane Collaboration, and ESRC - our task is to broaden this agenda to include concepts of 'evidence' and health, and appropriate means of assessing these.

The role of the Public Health Evidence Steering Group (PHESG) is to assist the HDA in developing appropriate and accessible ways of synthesizing and interpreting evidence for the effectiveness of interventions to improve health and reduce health inequalities. The Methodology Reference Group is specifically established, as a sub group of PHESG, to help the HDA consider the implications of the major questions that arise once a commitment is made to expand the potential scope and focus of the evidence base beyond data drawn from randomized controlled trials and similar methodologies. It is also the vehicle to draw together the work of the Cochrane Collaboration qualitative methodology group, the Cochrane Health Promotion Field, the Campbell Collaboration, the NHS Centre for Reviews and Dissemination, and others with respect to qualitative evidence.

Aims and objectives

Specifically, the Methodology Reference group is asked to consider the following issues:

- What frameworks exist that would allow us to include a broad range of data in our briefing documents?
- What are the epistemological positions on health and evidence inherent in these frameworks?
- How well do our current processes and tools meet the challenges outlined in this paper?

- Can we include data from process, qualitative, practice and action research within our current framework or do we need to develop a parallel protocol?
- How best might we begin to assess the effectiveness of interventions at reducing health inequalities?
- What are appropriate quality standards for evidence drawn from different disciplines and methodologies?

Outputs

The methodologies group will meet twice-yearly, and its purpose will be to provide an ongoing steer for the methodological and conceptual development of the HDAEB project. For this first year, we should like to suggest that the group to appoint one or two lead 'writers' and produce a review or document for open discussion on the issues above. However, the group is free to set its own method of working and terms of reference.

References

- Callinicos, A. (1999) *Social Theory: A Historical Introduction*, Cambridge: Polity.
- Cochrane, A. (1971) *Effectiveness and efficiency: random reflections on health services*. The Rock Carling Fellowship: British Medical Journal.
- Department of Health (1999) *Saving Lives: Our Healthier Nation*. London: Stationery office.
- Department of Health (2001) *A Research and Development Strategy for Public Health*. London
- Egger, M., Davey Smith, G., Altman, D.G. (2001) *Systematic Reviews in Health Care: Meta Analysis in Context*, London: BMJ Books.
- Foucault, M. (1973) *The Birth of the Clinic: An Archaeology of Medical Practice*. London: Tavistock.
- Giddens, A. (1982) *Profiles and Critiques in Social Theory*, London: Macmillan.
- Graham, H. (2000) (ed) *Understanding Health Inequalities*. Buckingham: Open University Press
- Habermas, J. (1987) *The Philosophical Discourse of Modernity*, Cambridge: Polity
- Health Education Board for Scotland (1996) How effective are effectiveness reviews? *Health Education Journal* 55: 359 - 362.
- Lupton, D. (1995) *The Imperative of Health: Public Health and the Regulated Body*, London: Sage.
- Mann, C. (1990) meta-analysis in the breech, *Science*, 249:476-80.
- Mays, N., Roberts, E., Popay, J. (2001) Synthesising research evidence. In Fulop, N., Allen, P., Clarke, A., Black, N. (eds) *Studying the Organisation and Delivery of Health Services: Research methods*. London: Routledge 188-220.
- Turner, B. (1992) *Regulating bodies: Essays in medical sociology*. London: Routledge

Annex 1: Draft search protocol

1. Introduction

This chapter details the process, which is undertaken to access information relating to the evidence base project from hard copy and electronic sources. It provides an overview of who is involved in the process, details of core literature searching strategies used, the range of databases scanned and the storage of literature once it is found.

2. Overall Aim of the Literature searching process

To carry out literature searches which will identify high quality systematic reviews, meta-analysis and other 'reviews' that provide information relevant to answering the following question.

What works to promote health and reduce health inequalities?

In 2001/2002 this question is being addressed across a range of topic areas including: the Under fives (Low birth weight, Social support in pregnancy and early motherhood, Breastfeeding), Health in later life (Falls, Hip fractures, Depression, Transport), Diet, Nutrition and Obesity (Nutrition in pregnancy, Breastfeeding, Obesity), HIV/AIDS/STIs, Inequalities (Generic search, plus Transport, Housing, Fiscal Policy (Income and Health)), Physical Activity, and Teenage Pregnancy topic groups.

3. Who is involved

The overall responsibility for developing this aspect of the evidence base project lies with the public health information team (PHIT), which includes a range of information specialists, officers and librarians. However, this team will work closely with research specialists at the HDA and a number of external agencies identified as having expertise in this area.

4. Working in Partnership

To date the research and information directorate is working in partnership with the:

- NHS Centre for Reviews and Dissemination (NHSCRD)
- Library and Information Development Unit (LIDU)
- Centre for Evidence-based Mental Health (CEBMH).

These agencies are actively supporting the evidence base project by carrying out a range of literature searches across the topic areas identified above. In addition the HDA will work with these agencies on the continual development and refinement of search strategies used as part of the evidence base project.

5. Step by Step Process

Detailed information about the process entailed in identify evidence based literature is described in appendix 2. Steps one to five provide a brief overview of the process as a whole.

1. *Developing core search strategies* - the PHIT work with the NHS CRD to develop and produce a range of core search strategies for use in searching electronic databases identified as relevant to the evidence base project.
2. *Developing topic specific search strategies* – research specialists with lead responsibility for particular topic areas work with the PHIT and the external agencies to identify topic specific and common terms for their areas. In so doing core strategies are adapted to maximize the number of relevant reviews retrieved by the search.
3. *Searching electronic databases* - the external agencies carry out the searches using a core set of electronic databases as identified in the protocol described in appendix 2. External agencies work with research specialists to identify any additional databases relevant to their topic areas.
4. *Storing searches on reference manager* – each of the external agencies will provide the evidence base team lead with a complete set of references arising from the search process. These will be uploaded onto reference manager for storage during the critical appraisal process.
5. *Transferring evidence base to externally available websites and databases* – once the critical appraisal process is complete the public health information team will liaise with research specialists to transfer details of studies from reference manager to the 2 principal databases used by the HDA to disseminate the evidence base project. These are HDAEB and Healthpromis. PHIT will transfer material onto the relevant databases according to how individual studies have be categorised during the appraisal process. The critical appraisal tool categorises studies as follows:
 - To be included as *data* where the whole of the review is judged to be of high quality (i.e. it forms part of the core material upon which we will base our evidence based statements in the Evidence Briefing Document).
 - To be included as data where only a part of the review is judged to be of high quality (Reviews judged to be 1 or 2 will automatically be included in the HDAEB Evidence Base Website whether or not it exists in a format that we can make available for downloading. Where it cannot be made available like this we will provide a link to another database [e.g. to DARE], or failing that simply a reference with an explanation.)
 - To be included in the Evidence Briefing Document as background or context. It may occasionally be appropriate to recommend that reviews in this section should also be made available on HDAEB.
 - To be included in only in Healthpromis database.
 - Discard

ANNEX 1 APPENDIX 1

HDA Evidence Base Core Search Strategy

The strategy below should be constructed for each of the topic areas under consideration.

The elements numbered 1-5 represent the different concepts for each part of the strategy, rather than the actual phrases or MeSH headings. Each element would encapsulate the range of appropriate MeSH headings. And, for each particular MeSH heading, we would also construct the equivalent freetext phrase with the appropriate use of connectors like 'adj'. This ensures that, regardless of available MeSH headings within each separate database, the search strategy remains comparable in scope.

1. General and specific terms to describe the topic/disease area, eg. HIV, obesity, drugs, ecstasy, alcohol, HIA, etc
2. Topic-specific prevention/interventions, eg. HIV prevention, food restriction, diet therapy, safer sex promotion, slimming clubs, condom provision, needle exchange, nicotine replacement, etc
3.
 - a) Specific health promotion/prevention activities/interventions, eg:
 - Policy/legal/economic/social policy
 - eg. Law on consensual sex, anti-discrimination re sexuality
 - Taxation
 - Sponsorship
 - Packaging
 - Licensing
 - Community safety
 - Promotion
 - Pricing
 - Changing attitudes
 - Support
 - Education – classroom, seminar, etc
 - Group work
 - Leaflet
 - Mass media/advertising/campaign

- Peer education
- One-to-one advice/counselling
- Outreach
- Harm minimisation
- Skills development
- Facilitation interventions – eg. training professionals
- Community and organisational development

b) Generic health promotion/public health terms, eg:

- public health
- health promotion
- behaviour therapy
- primary prevention
- preventative health services
- preventative medicine
- primary health care
- health education
- intervention
- behaviour modification
- early intervention education
- education
- health services
- community safety
- attitude change
- lifestyle change

4. Any particular Setting or Target Population for focus (if applicable), eg:

Populations

- Infants
- Children
- Children in care/looked after
- School excludes
- Persistent truants
- Children of drug using parents
- Young people
 - under 25s
 - 11-15
 - 16-18
 - 19-24
- Older people
- Women

- Pregnant women
- Ethnic groups
 - African
 - Zambia
 - Zaire
 - Kenya
 - Tanzania
 - Uganda
 - Zimbabwe
 - Asian
- Socio-economic groups
- Homeless
 - Homeless young people
- Men
 - Young men (hetero)
- Men who have sex with men
 - gay men
 - young gay men
 - non-gay-identified men
- People with (diagnosed) HIV infection/AIDS
- Sex workers/prostitutes
 - Male
 - Female
- Injection Drug Users
- Offenders
 - Young offenders
- Healthcare workers

Settings

- Family
- Education
 - Secondary School
 - University/college
- Workplace
- Health
 - Primary care (GP)
 - Hospital
 - Antenatal clinic
 - GUM clinic
 - Treatment centers/clinics (HIV/AIDS)
 - workplace for healthcare workers
- Needle exchanges
- Social Groups
- Pubs and Clubs etc

- Public Sex Environments (PSEs) – cruising sites, cottages, backrooms, saunas
- Helplines
- Internet – websites, discussion groups/ mailing lists
- Street
- Community (eg. gay community)
- Prisons

5. Exclusions, eg. developing countries, treatment, drug therapy, Africa, etc.

6. Review Strategy for the database in question

7. Limits: 1996 to date; human; English language

The search strategy is connected thus:

((2 OR ((3a OR 3b) AND 1)) AND 4 NOT 5) AND 6) AND 7

Once this search has been completed, a sub-search for references related to inequalities should be performed, eg. inequalities, low income, minority, etc.

ANNEX 1 APPENDIX 2 Working with commissioned Agencies

The HDA has commissioned 3 agencies (NHS CRD, CEBMH and LIDU) to support the literature searching aspects of its evidence base project.

The following extract from the agreements set up between the HDA and the agencies details the process entailed in accessing evidence-based material.

Extract:

Specifically you will - working in conjunction with the HDA's Research and Information Directorate:

1. Attend a pre-meeting at the HDA to meet others working on the literature searching task and members of the HDA's Research and Information Directorate to clarify roles and objectives.
2. Work with the other organisations involved in carrying out literature searches for the HDA's Evidence Base, you will develop consistent search strategies to identify high quality systematic reviews, meta-analysis and other 'reviews'. Building on the work already undertaken by the HDA (we have used the York CRD strategy 1 for Medline and the York CRD strategy for Cinahl and Health Evidence Bulletins Wales strategies for Embase and PsycINFO), you will help develop review search strategies for databases, where no published review strategies exist. You are expected to attend the pre-meeting identified above and share search strategies via e-mail with others involved in this process. It is anticipated that the resulting search strategies will be published on the Public Health electronic Library.
3. With HDA's researchers identify topic specific search terms. See also 6 below.
4. In the first instance you will be expected to carry out a review of databases of reviews and guidelines and then in liaison with the topic teams, broaden your search to include the major bibliographic databases for the subject area. These searches will also consider inequalities in health, community development and health impact assessment (See 6 below).

An indicative list of databases for this task is given below:

Completed and Ongoing Reviews
Cochrane Library: Cochrane Database of Systematic Reviews
DARE < http://www.york.ac.uk/inst/crd >
National Research Register (including CRD Ongoing Reviews) < http://www.update-software.com/National/ >
Health Technology Assessment Database < http://www.york.ac.uk/inst/crd >
SIGN Guidelines http://www.sign.ac.uk
National Guideline Clearinghouse http://www.ahcpr.gov/clinic/assess.htm
National Coordinating Centre for Health Technology Assessment < http://www.hta.nhsweb.nhs.uk >
NICE web page (published appraisals) < http://www.nice.org.uk/nice-web/ >
HSTAT < http://text.nlm.nih.gov/ >
The DH Research Findings electronic Register < http://www.doh.gov.uk/research/rd3/information/findings.htm >
Indexes to and summaries of Clinical Effectiveness Sources Including Reviews, Appraisals of Reviews, and Evidence Based Guidelines.
TRIP < http://www.tripdatabase.com >
SchARR Lock's Guide to the Evidence < http://www.shef.ac.uk/uni/academic/R-Z/scharr/ir/scebm.html >
Clinical Evidence: A compendium of the best available evidence for effective health care. Latest edition. Reference collection
Health Evidence Bulletins Wales < http://www.uwcm.ac.uk/uwcm/lb/pep >
Searches for ongoing trials - to be conducted by agreement with reviewers.
CCTR (Cochrane Library)
Controlled Clinical Trials (< http://controlled-trials.com >)

List of systematic and other databases suggested by Julie Glanville, NHS CRD York, personal communication, (2001).

- Having exhausted the databases outlined above, and as part of the process of developing the search strategy, you will - working with the HDA's Research and Information topic teams, suggest appropriate additional bibliographic databases to search and carry out these searches. At a minimum it is expected you will search Medline, CINAHL, PsycINFO, EMBASE, Sociological Abstracts; and Evidence Based Medicine Reviews, although you will be expected to recommend appropriate databases. This will be carried out in consultation with the HDA.

Note: We are also interested in review articles and not exclusively systematic reviews and meta-analyses. Where no systematic reviews, meta-analyses or other review articles exist you should consult with the HDA before seeking to identify primary studies for the topic areas.

- For both 4 and 5 above, you will make use of the HDA Evidence Base search strategy (attached) and carry out additional searches for each topic area and include terms for health impact assessment, community development and inequalities (attached). Suitable terms will need to be developed for health impact assessment and additional terms can be suggested for inequalities and community development. Suggestions as to how these terms may be improved and incorporated in the searches are welcome and should be fed into the e-mail groups.
- Liaise (e-mail discussion groups, meetings, telephone for 1. the overall search process and 2. each topic group) with others working on the literature searching task to ensure

consistency across the search strategies and refine searches in accordance with the HDA's wishes.

8. Provide output suitable for uploading into the Reference Manager bibliographic database software.
9. Obtain articles as requested by the topic teams at an additional charge per item to be agreed with the HDA.
10. Attend a post-search process meeting at the HDA to share experiences of this process and suggest how possible future exercises could be improved.

ANNEX 2 Critical Appraisal Tool

Health Development Agency Critical Appraisal Tool for Evidence Base

(17/12/01 version)

Introduction

The Critical Appraisal Form is designed to assist judgements about the quality of review material, which will form the data for the Evidence Base. The tool should be read and used in conjunction with the document *Protocol for Selecting Reviews for Evidence base*. The principal purpose of the tool is categorize reviews into one of five types. These are:

- To be included as *data* where the whole of the review is judged to be of high quality (i.e. it forms part of the core material upon which we will base our evidence based statements in the Evidence Briefing Document).
- To be included as data where only a part of the review is judged to be of high quality (Reviews judged to be 1 or 2 will automatically be included in the HDAEB Evidence Base Website whether or not it exists in a format that we can make available for downloading. Where it cannot be made available like this we will provide a link to another database [e.g. to DARE], or failing that simply a reference with an explanation.)
- To be included in the Evidence Briefing Document as background or context. It may occasionally be appropriate to recommend that reviews in this section should also be made available on HDAEB.
- To be included in only in Healthpromis database.
- Discard

Normally we would expect that papers judged as 1,2 and 3 would also be included in Heathpromis.

The secondary purpose of the tool is to assist the team in identifying key elements for the writing of the brief. Several sections in the Critical Appraisal Tool seek details, which will be useful at subsequent stages of the process. Obviously it may not be possible to identify answers to all the questions in the tool in every

paper read, and that is not the expectation. However, it will probably be the case that the more of the questions that can be completed, the easier the judgement will be about allocation to the categories. Where it is not possible to find answers to the questions in an initial reading of the paper, judgements are likely to be more difficult. Given that the business is making judgements, the fact that they may be difficult is simply part of the task itself. However, the processes need to be transparent so others can see easily how the judgements were reached and indeed to observe how, on occasions, it was necessary to work with relatively thin evidence.

There are questions about bias in the Critical Appraisal Tool. In systematic review, overcoming bias in order to make sound judgements' is of paramount importance. Our task when allocating reviews to the categories above and then in writing the briefing documents is, amongst other things, to ascertain how well the authors have made attempts to overcome the biases. There are two types of bias. Those inherent in the primary research the authors have considered and those, which they may have introduced themselves in the process of doing the review. The Critical Appraisal form does not allow us to make judgements about biases in the primary work. However, it should be possible to see how well the authors have dealt with, or are aware of biases in the primary work. The tool allows the user to scan for several 'give away' features. These include double counting (where the results of the same study have been reported in different papers, and those papers are treated as separate examples in support of the evidence); self references (where the authors use their own work to add strength to evidence accumulated); self interest (where the reviewers have a particular stake in the spin on the evidence); foreign language studies not included (there is a known systematic publication bias among non- first language English speaking scientists, to publish positive findings in the English Language literature).

The best source on bias is Egger, M., Davey Smith, G, Altman, B.G. (2001) *Systematic Reviews in Health Care: Meta-analysis in Context*, London: BMJ Books.

Author(s): _____

Title: _____

Full Bibliographic Details

List the topic areas with which the review is concerned

Does this paper address your topic/research question?

Is the paper best described as a: (tick as appropriate)

- Systematic review?
- Meta-analyses?
- Syntheses?
- Literature review?
- Other review? (please specify)

Does it address (tick as appropriate):

- Effectiveness (interventions and treatments)
- Causation
- Monitoring & surveillance trends
- Cost
- Inequalities
- Other (please specify)

Is it worth continuing?

<p>Does the review have a clearly focused aim or research question? Consider whether the following are discussed:</p> <ul style="list-style-type: none"> • The population studied • The interventions given • The outcomes considered • Inequalities <p>If the research question or aim <i>is</i> clear, does the review allow a judgement to be made about:</p> <ul style="list-style-type: none"> • The strengths of the evidence? • The weaknesses in the evidence? • The gaps in the evidence? • The currency of the evidence? <p>Do the authors make judgements about the quality or strength of the evidence or do they attempt to rank or calibrate it in some way?</p>	<p>Yes</p> <p>Yes</p> <p>Yes</p> <p>Yes</p> <p>Yes</p> <p>Yes</p> <p>Yes</p> <p>Yes</p> <p>Yes</p> <p>Yes</p> <p>Yes</p>	<p>No</p> <p>No</p> <p>No</p> <p>No</p> <p>No</p> <p>No</p> <p>No</p> <p>No</p> <p>No</p> <p>No</p> <p>No</p> <p>No</p>	<p>Unsure</p> <p>Unsure</p> <p>Unsure</p> <p>Unsure</p> <p>Unsure</p> <p>Unsure</p> <p>Unsure</p> <p>Unsure</p> <p>Unsure</p> <p>Unsure</p> <p>Unsure</p>
<p>If the evidence is calibrated, ranked or categorised, what measure/scale is used?</p> <hr/> <hr/>			
<p>Is it worth continuing?</p>			
<p>Do the authors acknowledge theoretical issues in:</p> <ul style="list-style-type: none"> • The materials they have reviewed? • Their own approach? 	<p>Yes</p> <p>Yes</p>	<p>No</p> <p>No</p>	<p>Unsure</p> <p>Unsure</p>
<p>What measures of social difference do the authors use? (e.g. Class, occupation, socio-economic group, gender, ethnicity, age, residence, geography, disability).</p> <hr/> <hr/> <hr/>			

Do the reviewers try to identify all relevant English language studies	Yes	No	Unsure
Consider whether details are given of:			
• Databases searched	Yes	No	Unsure
• Years searched	Yes	No	Unsure
• References followed up	Yes	No	Unsure
• Experts consulted	Yes	No	Unsure
• Grey literature searched	Yes	No	Unsure
• Search terms specified	Yes	No	Unsure
• Inclusion criteria described	Yes	No	Unsure
• Sensitivity and specificity	Yes	No	Unsure
• What materials were excluded	Yes	No	Unsure
• Whether the data extraction was performed in a systematic way	Yes	No	Unsure
• Whether the criteria used to assess the quality of the primary studies was stated	Yes	No	Unsure
Is the primary source used by the reviewers drawn from:	Yes	No	Unsure
• Peer reviewed published materials	Yes	No	Unsure
• Non peer reviewed published materials	Yes	No	Unsure
• Unpublished materials	Yes	No	Unsure
• Self referential materials	Yes	No	Unsure
Do the reviewers consider non English Language primary sources?	Yes	No	Unsure
Is it worth continuing?			

Does the review make clear what steps have been taken to deal with potential bias?	Yes	No	Unsure
If yes, what are these?			

Have the authors taken care to avoid double counting of primary data?	Yes	No	Unsure
Do the authors refer to primary research studies in which they themselves have been involved?	Yes	No	Unsure
Do the authors have a vested interest in the direction of the evidence?	Yes	No	Unsure
<p>If bias has not been overtly considered, or only partly considered, what are the potential biases, which should have been acknowledged?</p> <hr/> <hr/> <hr/>			
<p>To what extent does the treatment of bias in the in the paper affect any conclusions in it, about strengths of evidence?</p> <hr/> <hr/> <hr/>			
Do the authors address the quality (rigour) of the included studies? Consider whether the following are used:	Yes	No	Unsure
<ul style="list-style-type: none"> • A rating system • More than one assessor 	Yes Yes	No No	Unsure Unsure
If study results have been combined, was it reasonable to do so? Consider the following:	Yes	No	Unsure
<ul style="list-style-type: none"> • Are the results of included studies clearly displayed? • Are studies addressing similar research questions? • Are the studies sufficiently similar in design? • Are the results similar from study to study (test of heterogeneity)? • Are the reasons for any variation in the results discussed? 	Yes Yes Yes Yes Yes	No No No No No	Unsure Unsure Unsure Unsure Unsure
Have the data been presented in a way which allows an independent assessment of the strength of the evidence to be made?	Yes	No	Unsure
Can statements made by the reviewers be tracked back to primary sources precisely (by page reference)?	Yes	No	Unsure
	Yes	No	Unsure

What is the overall finding of the review? Consider:

- How the results are expressed (numeric – relative risks, etc.)?
- Whether the results could be due to chance (p-values and confidence intervals)?

Do the authors acknowledge any weaknesses in what they have written?

Are sufficient data from individual studies included to mediate between data and interpretation/conclusions?	Yes	No	Unsure
Does this paper cover all appropriate interventions and approaches for this field (within the aims of the study)?	Yes	No	Unsure

If no, what?

Relevance			
Can the results be applied/ are generalisable to a UK population/ population group?	Yes	No	Unsure
• Are there cultural differences from the UK?	Yes	No	Unsure
• Are there differences in health care provision with the UK?	Yes	No	Unsure
• Is the paper focused on a particular target group (age, sex, population sub-group etc)?	Yes	No	Unsure

Recommended category 1, 2, 3, 4 or discard.

Additional comments:

Reviewer: _____

Date

ANNEX 3 Critical Appraisal Tool Protocol

Protocol for selecting reviews for Evidence Base and for Evidence Briefing Documents

Introduction

The purpose of this paper is to outline the selection criteria for inclusion of reviews in evidence base, the web-based database on the HDA-EB website. It also describes the relationship of HDA-EB and EB to the parallel bibliographic resource healthpromis, and the umbrella public health information site PheL (public health electronic library).

Types of Evidence

The Evidence Based Briefing Documents will be constructed by distinguishing between two types of review material: that upon which Evidence Based statements will be based; and that which may be used for background and context for the briefing document. The former provides *data*, which will be interrogated to answer a series of questions. The series of questions relate to strengths and weaknesses, gaps, currency, cost effectiveness, and relationship to reductions in inequalities. For some of the topics with which we will deal, *data* will be disproportionately drawn from Systematic Review, while contextual commentary will be drawn from narrative reviews.

Not everything called Systematic Review will meet the criteria as data for inclusion in the Evidence Based Briefing Documents. We will have to attend to problems of bias, and the degree of rigour of analytic technique in review material, in order to be satisfied that it meets the criteria for inclusion as data. The dividing line between systematic and narrative reviews is clear-cut in theory, but in practice may not be easy to apply. There are poor systematic reviews (and we have already seen examples of this in the Cochrane collection), and not all narrative reviews are either unsystematic nor lacking in useful data. In some of the topics we will be working on there may only be narrative material. We have to be systematic *and* flexible.

Threshold Rules

Rather than devise a set of criteria to distinguish between Systematic Reviews and other types of review, our method will be to apply clearly defined threshold rules to what we read and use the outcome of the application of the rules to determine whether and where the material is included. We will need to recognize that the threshold rules will not make the decision for us, but we will make judgements, using the rules to help. It will sometimes be difficult to make the decisions and we will therefore need to be clear about the processes we use, and document them carefully, so it is open to audit and challenge. The need for

transparency and openness means that we must retain all the records of our decision-making processes. We will not always reach a consensus, and we will therefore need to acknowledge that differences of opinion are legitimate and normal. We also need to recognize that our criteria will evolve and develop first as we use them, and second when the Public Health Evidence Steering Group begins to offer advice on applying our criteria much more broadly to materials which have not traditionally found their way into any kind of review. In some areas the data will be much stronger than in others.

The process

The procedure for the selection of materials will be as follows.

- Receipt of long list of papers generated by the search strategies.
- Abstracts read by topic lead, and potential reviews separated from long list for detailed consideration (the short list), marginal abstracts identified for consideration by second reader.
- Second reader consults marginal abstracts and identifies further reviews for the short list.
- Lead and second (and third) readers divide up the short list.
- Each reader assesses the short listed reviews using the critical appraisal tool. Two readers working blind read each review.
- Each review is allocated to one of five categories. These are:
 1. To be included as *data* where the whole of the review is judged to be of high quality (i.e. it forms part of the core material upon which we will base our evidence based statements in the Evidence Briefing Document);
 2. To be included as data where only a part of the review is judged to be of high quality (Reviews judged to be 1 or 2 will automatically be included in the HDAEB Evidence Base Website whether or not it exists in a format that we can make available for downloading. Where it cannot be made available like this we will provide a link to another database [e.g. to DARE], or failing that simply a reference with an explanation.)
 3. To be included in the Evidence Briefing Document as background or context. It may occasionally be appropriate to recommend that reviews in this section should also be made available on HDAEB.

4. To be included in only in Healthpromis database.
5. Discard

Normally we would expect that papers judged as 1,2 and 3 would also be included in Heathpromis.

The Critical Appraisal Tool

Presently the critical appraisal tool helps us to discriminate between 1,2,3, or 4. In order to do this the critical appraisal tool tackles a number of questions.

- If the research question that the review deals with is clear, does the review allow us to judge how strong the evidence marshaled in the paper is?
- Is strength of the evidence calibrated in some way, and if so what is the calibration?
- Does the review make plain that steps have been taken in the course of aggregating results or data, to reduce bias? What are those steps?
- If bias has been only been partly addressed or not considered at all, what are the potential bias(es) which should have been included and acknowledged by the original writers of the reviews?
- To what extent does the treatment of bias in the paper affect its conclusions about the strength of the evidence?
- Have the data on which the paper is based, been provided in a full enough way for readers of the review to judge for themselves, independently, the strength of the evidence?
- To what extent does the review consider the theories of change, which may underpin their assessment of the evidence?
- Have the authors made plain what their search strategies were?
- Is it possible to replicate the search strategies?
- Is it possible to identify search strategies that the reviewers have not conducted?
- Can the statements made by the reviewers about the evidence be tracked back to primary sources exactly i.e. to specific pages in specific source documents?

- Is the source material trustworthy? This is difficult to judge and two conflicting assessments need to be made. Is it drawn from reputable peer reviewed sources? If it is not drawn from reputable peer reviewed sources why is this? Is it because the reviewers have accessed unpublished data in order to overcome publication bias? Is the source grey literature? We must be very cautious about lists of references which are self referential, and which refer to derivative sources such as fact sheets produced by pressure groups. Is the apparent strength of the evidence an artifact of double counting of studies? Does the strength of the evidence depend on the use of fact sheets from pressure groups and or from self-referred work?
- Are the authors explicit about any analytic processes they have used such as meta analysis or other technique? How familiar are we with the analytical technique, which is being used?
- Do the reviewers have a vested interest in the direction of the evidence?
- Do the authors of the review acknowledge the weaknesses in what they have written?

We could turn the critical appraisal tool into a scoring tool. However at this stage it is uncertain what weights could be given to the different dimensions identified. The questions in the critical appraisal tool are best thought of as articulations of the kinds of thing that is usually implicit in the judgements made of reviews, we are simply making them explicit. It may be possible that with use, we will be able to do develop a scoring system. We are some way off from that at the moment.

The judgement that will need to be made in other words is precisely that, a judgement. The process that we should engage in is to read the paper to assess it using the criteria, and then on the weight of the evidence of the reading, judge it to be category 1,2,3 or 4. Reading blind, the second reader will make an independent judgement. There should be a meeting to decide where to allocate the review. Where there is agreement about the categorization, the papers can be logged for the writer of the briefing document to start work on. Where there is disagreement, each reader should explain to their partner why they reached the judgement they did. No attempt to *persuade* other should be made. On the basis of comparing the assessments made using the different dimensions of the critical appraisal tool, a reader may change their judgement. Where differences are unresolved. Another reader should be consulted if the topic lead thinks it necessary. The third reader will make their judgement independently and although they will know that there was disagreement, they should not be told what the disagreement was. The decision about category will be made by simple majority.

The writer then may begin to construct the evidence briefing document. It is entirely likely that once the writing begins, that some of the judgements about quality will be refined, and the lead author may decide that a particular contribution should be reevaluated. The final responsibility for where material rests will remain with the lead writer. Only when the document is completed should the technical team start adding materials to databases and then only according to the contents of the document not to the critical appraisal tool. Obviously if the lead writer makes changes to judgements these too must be recorded on the critical appraisal tool along with an explanation so that both we and anyone challenging our judgements can see the process very clearly.