A. Aims and research questions

We aim to revolutionise the way in which evidence about behaviour change interventions (BCIs) is synthesised and used to advance our understanding of behaviour and behaviour change. We will build and apply an Artificial Intelligence (AI) System to answer the question: **'What works, how well, for whom, in what setting, for what behaviours - and why?'.** We will generate major advances in (i) behavioural science theory and practice, and (ii) computer and information science.

Our proposal focuses on the development of an AI System that uses Natural Language Processing (NLP) and Machine Learning to extract information from the burgeoning scientific literature, reasons with it to create new understandings and continually learns from new information fed into it to improve its performance. We will also build and test a sophisticated online user interface to interact with this AI system.

The behavioural science challenge is to develop (i) a BCI 'Ontology' (i.e. a formal structure for organising knowledge in the form of elements and relationships between these) to represent key information that needs to be encoded from a diverse scientific literature, (ii) a method of training an NLP system in order to find and extract this information in published research reports (journal articles), and (iii) a user-friendly interface that explains its conclusions in a way that people can understand and trust.

The computer and information science challenge is to turn the BCI Ontology into an AI system involving a set of computerised knowledge structures, an NLP system that can be taught to extract the information needed from research reports, and a computerised reasoning system that can enhance the knowledge structures to answer the wide variety of questions that may be asked and learn from the feedback it receives.

The key **scientific questions** are:

1. Behavioural science

a. What are the key features of BCIs, usage patterns, contexts (populations and settings), mechanisms of action and targeted behaviours that need to be captured to build a BCI Ontology?
b. How can this Ontology be used to inform an AI System for understanding and predicting behaviour change?
c. What are the key features of a user interface to interact with the AI System to make it easy to use and provide answers that are understood and trusted?
d. What is the added value of the AI System over traditional methods of evidence synthesis?

2. Computer and information science

a. What are the effective knowledge structures needed to represent key features of BCI evaluations and the interrelations between these features?
b. What kinds of NLP system can be developed to learn to extract relevant information from research reports in order to populate those knowledge structures?
c. What kind of Reasoning System can be applied to draw inferences from the information in the BCI Ontology to enhance the knowledge structures?
d. What kinds of Machine Learning algorithms can be developed to improve the accuracy of the NLP system, the knowledge structures and the Reasoning System as new information about the interventions, contexts and target behaviours becomes available?

3. Interdisciplinary science

a. What new insights about BC are generated by the AI system interacting with BC experts?
b. How can behavioural, computer and information scientists work most effectively together to create a large scale AI System?

B. Essential background

The major challenges facing society (health, environmental, climatic etc.) require behaviour to be changed. For example, obesity, antimicrobial resistance and infection transmission can be mitigated by (respectively) healthy eating, appropriate antibiotic prescribing and vaccination uptake. The recent Ebola outbreak emphasises the need to understand health problems behaviourally and respond with effective interventions, which may be environmental, person-delivered and/or technological, to change behaviour.

Over 2000 evaluations of BCIs are published each day; however, the majority of research is wasted in terms of its contribution to evidence and practice, due to reporting in unstructured and/or heterogeneous formats with terminology that is not standardised.[1] Without a systematic and shared way of structuring the evidence, efforts to replicate studies, implement effective interventions and accumulate knowledge are severely curtailed. As a result, despite the accelerating production of evidence, major behavioural problems are approached using interventions based on partial, fragmented evidence, with often limited success, with huge wasted resource and opportunity.[2]

While developing a science of BC has been an active area for many decades, surprisingly, there has been no concerted effort to organise and analyse the large volume of intervention evidence that is generated taking into account all the key factors that are relevant. Methods such as meta-analysis and meta-regression have been used to combine evidence to answer specific questions as to whether an intervention works. However, they are not able to take sufficient account of the numerous contextual factors that influence intervention effectiveness; neither are they able to reason from that information to new interventions, contexts and behaviours.

Our research programme aims to transform this situation completely by developing and evaluating a shared organising framework which relates intervention features (and combinations of features in different contexts) with how they achieve their intended effects. We will apply this BCI Ontology at scale to systematise the knowledge implicit in BCI research reports. An ontology is an information structure for codifying collective knowledge so that the information can be digitalised.[3] (Note: 'Ontology' has a different meaning on philosophy).

The BCI Ontology will revolutionise the field by organising the world literature in a way that can address continuously how BCIs work and why they vary in effectiveness, enabling the efficient development of more effective and context-appropriate interventions. Linking this with an open-source, accessible interface will be invaluable to scientists, practitioners and policy-makers globally.

SM leads research that has created a paradigm shift in the field of behavioural science by developing methods for specifying the content of interventions, i.e. taxonomies of BC techniques.[4,5] Our very recent key papers developing these methods of integrating BCI evidence have already been cited >2500 times, and the methods developed thus far are being used by leading organisations in the UK such as NICE and Public Health England, and in the USA by the National Institutes of Health. Over 3000 researchers, practitioners and policy makers have been trained in their use by SM and colleagues. An independent publication by 11 co-authors from 12 Centres, spanning research and policy, observed: "*Within behavioral science, it is heartening to see efforts such as Michie and colleagues who have developed a taxonomy of behaviour change terms to help create evidence-based guidelines for implementation research. It is hoped that these types of taxonomy and ontology development efforts ... will eventually lead to faster, more efficient scientific discoveries that will lead to better and more sustainable behaviour change*".[6]

The Director of NIH's Office for Behavioral and Social Science Research considers this work to be a global priority, stating in his letter of support: "*I've been particularly interested in your work and your efforts to bring order to the cacophony of BC techniques in the literature… Although the NIH is supportive of all aspects of this project, we are particularly interested in supporting the international aspects of this work and the greater dissemination of these approaches in the United States, so travel and logistic support for some of the expert network and international consortium activities seems an appropriate use of the supplemental funds that NIH is willing to consider providing should this grant be funded.*"

To accelerate this progress we will combine (i) behavioural science to clarify terminology and build the BC framework with (ii) computer and information science to address the overwhelming quantity of evidence and apply reasoning processes at a scale that cannot be achieved by humans. The proposed research will extend our work to date to develop taxonomies (part of the BCI Ontology) of behavioural outcome, modes of intervention delivery and mechanisms of action thus laying the foundations for a BCI Ontology. This has the potential to revolutionise the field, but behavioural science alone will never be able to keep up with the accelerating rate of evidence production: collaboration with advanced computational methods is needed.

Until recently the proposed project would not have been feasible. However, new developments in AI have transformed the picture. In particular, we will use the suite of AI applications known collectively as IBM Watson, developed to optimise use of the large, accumulating evidence bases. It has been successfully applied in areas of medicine[7,8,9] and banking[10], among others, where it has facilitated human decision-making. For example, it has been used to interrogate a corpus of the oncology literature and recommend treatment options by interpreting the patient's profile.[11] 'Watson Oncology' also includes new algorithms developed by IBM Research to extract risk information from the research literature to help doctors to better use the collective knowledge about factors and disease when making decisions.[12] For more details about IBM Watson, see http://www.ibm.com/smarterplanet/us/en/ibmwatson/what-is-watson.html. IBM has also developed reasoning systems to explain how decisions were derived in a range of projects, which we will draw on for this project.[13,14,15] Watson and these reasoning systems will not in themselves be enough but the technology will be foundational to the proposed project.

The BCI Ontology will provide a framework to guide NLP algorithms adapted from the Watson system to read and organise the thousands of individual reports of research evidence that make up the BC field. Building on the evidence base and the BCI Ontology, the AI system will identify key features and relationships between intervention content, delivery and contexts (populations and settings), mechanisms of action, and specific behaviours. It will assimilate data at a pace and scale that matches the rate of new publications, beyond human capacity. It will also identify hitherto un-investigated intervention strategies, and suggest new causal explanations.

This system will be an invaluable tool for policymakers and practitioners, offering, for the first time, a comprehensive and systematically organised repository of real-time evidence of effective interventions for specified real-world settings. The current slow rate of advance in BC will thereby be replaced by accelerating scientific developments and applications. Making this system accessible and user-friendly for decision-makers is a key objective of this project; equally important is the feedback that the system will get from being used and critiqued for use in real-world settings.

The project team will develop, evaluate and apply both the NLP algorithms and the Reasoning System both of which will draw heavily on Machine Learning. Key steps are (i) development of an automated system for extracting key features from research reports, (ii) development of knowledge structures from the BCI Ontology and (iii) development of a Reasoning System to make interpretations and predictions. The basis of our scientific advance for computing within our collaboration will be to extend the existing functionality of computerised modelling to learn from its interaction with the research literature and BC experts, thus providing new insights about BC (the Annexe outlines the technical specification of this work).

C. Approach

As with biomedical ontologies (e.g., the Human Genome Project), the BCI Ontology will reduce concept ambiguity and knowledge fragmentation by formally and transparently organising knowledge. The Ontology will be developed iteratively, with interactions between BC experts and Machine Learning algorithms informing the knowledge structures. The resulting AI System will be able to characterise existing research and facilitate the discovery of hitherto hidden knowledge about the causal mechanisms underlying BC. When the sum of our BC knowledge has been incorporated within this unifying system, we will be able to continuously update real-time evidence and enable scientists to test and develop theories of BC and rapidly access up-to-date evidence regarding optimal interventions and delivery for specific behaviours, populations and settings.

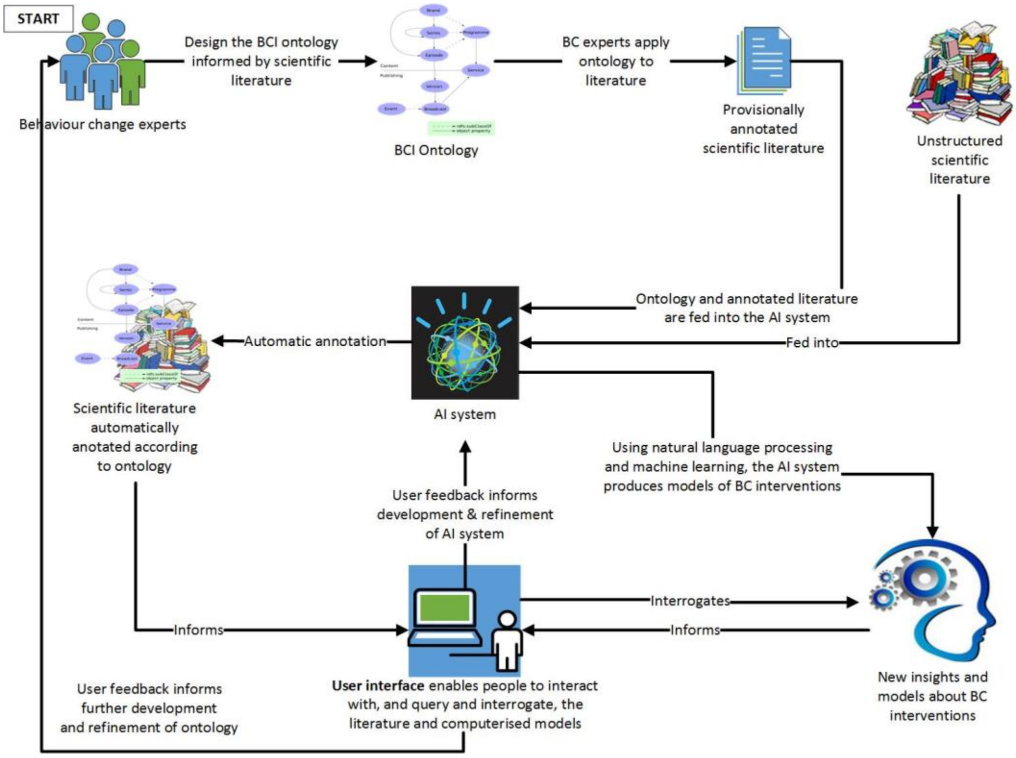
BC evidence is currently gleaned from unstructured texts through a largely manual process of analysing heterogeneous information which is lengthy, painstaking and error-prone. Findings and outcomes may be contradictory from one intervention to another, leading to incomplete understanding or inappropriate interpretation; and because the evidence synthesis process is slow, the evidence is often well out of date by the time of completion. While intelligent systems have been used to identify relevant research reports[16], their potential for data extraction and interpretation has not been exploited, although there is good evidence for this method in the medical informatics literature.[17] AI Systems, such as Watson, can revolutionise this process by:

1. extracting evidence using NLP and organising it by structuring, connecting and contextualising data according to the BCI Ontology and
2. interpreting the evidence by generating explanations and predictions.

These systems can transform unstructured evidence into recommended interventions and associated explanations for specific conditions, ranked by confidence with regard to the effectiveness of interventions in a given context.

This collaboration between behavioural, computer and information scientists will be able to show how current thinking about BC is organised in the world literature and will allow researchers and others to access and make sense of global evidence quickly. This will be achieved in three parallel interlinking Work Packages, bringing together in iterative fashion: BC experts and users, published evidence and AI. The information and knowledge flow is illustrated in Figure 1. Success will be evaluated by comparison with traditional evidence syntheses on cost-effectiveness and quality of outputs as rated by experts blinded to the source (see WP3).

Figure 1: Information and knowledge flow in the project



WP1: Organising BC Knowledge: Building a BCI Ontology, a corpus of annotated research reports and an international consortium

WP1 will build a **coherent** and **comprehensive** ontology for organising features of BCIs and their evaluations. This in itself represents significant progress for behavioural science. We will use this framework to annotate a large body of published BCI evaluation research, involving the international scientific community. Community buy-in for our research is key to gaining acceptance and for overall quality control, and has informed our thinking throughout this proposal.

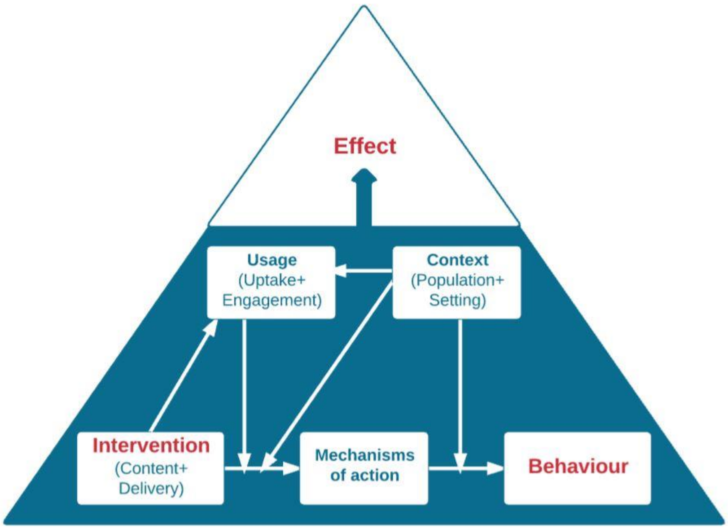
This **BCI Ontology** will, for the first time, characterise and link the following five key components of BCIs in a comprehensive and coherent way (Figure 2).

i. **target behaviours** comprising type of behaviour and specific outcome measures
ii. **contexts** comprising target population, setting and other relevant behaviours
iii. **interventions** comprising 'content techniques' and 'delivery'
iv. **usage** of the intervention, comprising uptake and engagement

v. **mechanisms of action** comprising the processes by which interventions have effects on behaviours.

This intervention-behaviour complex can then be linked to effect sizes in empirical evaluations.

Figure 2: Top level of the BCI Ontology



The top-level constructs of the BCI Ontology that the project will start with are shown in Figure 2. It depicts, starting from the bottom left, that characteristics of **interventions** (i.e. their content and delivery) lead to changes in **behaviour** through designated **mechanisms of action**. For example, providing ultrasound feedback on carotid artery stenosis may increase the likelihood of a smoker making a quit attempt by raising anxiety about the personal risks from smoking. Intervention **usage** (uptake and engagement) may be affected by the intervention and influence the impact of the intervention on the mechanisms of action. For example, some smokers may be reluctant to subject themselves to ultrasound feedback leading to low uptake, and/or may disengage emotionally from it when they start to hear unwanted information. **Context** (target population and setting) may influence usage as well as the impact of the message on the mechanisms of action and the impact of mechanisms of action on the behaviour. For example, heavily nicotine dependent smokers may be less likely to expose themselves to the information. If they are exposed to it, they may block out the messaging to reduce their anxiety levels. Even if they feel anxious, they may be less likely to try to stop smoking because of low confidence that they can succeed. Thus the intervention effect in a given population or for a given individual may vary as a function of the key moderating variables of usage, and context.

Each top level component of the Ontology will contain hierarchically organised sub-components with labels and definitions. These will be used to annotate the research reports for training and testing the AI System (WP2).

WP1 comprises two key tasks:

1. Build the Behaviour Change Intervention Ontology

BC experts (recruited from our international networks of >1000 experts) will work with us to build and refine the Ontology, starting with our preliminary Ontology (see table) based on our published and current work. It consists of hierarchical taxonomies of reliably annotated 93 BC techniques, 159 modes of delivery, 107 mechanisms of action and 256 behaviours.

Examples are provided below:

Key components	**Preliminary Ontology contains**	**Examples for smoking**	**Examples for physical activity**
Target behaviour	256 behaviours	Quit attempts/ quit success for >6mths	10,000 steps per day; Reduced sedentary time
Intervention: 'content' (techniques) and 'delivery'	93 behaviour change techniques 159 modes of delivery	Content techniques: goal setting, self-monitoring, restructuring the physical environment; Delivery: face-to face; classroom; a smartphone app	Content techniques: self-monitoring, prompts and cues, practical social support; Delivery: automated reminders; a smartphone app; exercise class
Context: target population and setting		Low income populations; people with mental health problems	Obese adults; people with Type 2 diabetes; Depressed teenagers
Use of the intervention in practice (uptake and engagement)		Use of Stop Smoking Services: quit attempts and attendance	Uptake of smartphone apps and continued engagement with these
Mechanisms of action: the processes by which the intervention affects behaviour	107 mechanisms	Craving reduction by pharmacotherapy	Improved action-planning

To refine this preliminary ontology, create content and extend it to other BCI Ontology components (e.g., population, setting and usage), we will conduct expert consensus exercises using online Delphi techniques[18] and 'webinars', as used effectively in our previous research.[5,19] These will ensure discrete, non-redundant elements, identify possible additional ones and ensure sufficient clarity of labels and definitions for satisfactory inter-rater reliability. The Ontology will be developed using a shareable, hierarchically-organised digital database in WebProtege, an open-source editor and framework for building intelligent systems (http://protege.stanford.edu/).

BC experts will work iteratively with the AI System created in WP2, generating potentially matching synonyms, concepts and terminology which will be verified by BC experts to achieve standardisation. The AI System will also specify relationships between components which will be tested using literature reviews and expert consensus methods. In this way, the AI System will propose new classificatory and organisational principles, the Ontology will become populated with evidence and the amount of expert effort needed to classify new research reports according to the Ontology will decrease.

2. Prepare BCI Ontology-annotated research literature

To 'train' the NLP part of the AI System (see WP2) in correctly extracting key features from research reports, we will use a corpus of 300+ published quantitative evaluations of BCIs that we have already annotated (coded and attached to the relevant text) using our preliminary Ontology. To 'test' this system we will select an additional corpus of research reports and annotate using our refined Ontology. Testing will start with a domain with simple behaviours, robust outcome measures and relatively coherent evidence (smoking) and then move to domains with more diverse behaviours and outcomes and heterogeneous evidence (alcohol, diet, physical activity). We will use reports published since 2010 in the first instance, as scientific and reporting standards have improved over time.

WP1 will produce a comprehensive, annotated evidence base of published evaluations of effectiveness of BCIs, annotated according to intervention features for use by WP2. It will then be maintained as an updating database of current knowledge, through automated searching, indexing, construct and effect size identification, comparison and classification, using technology described in WP2.

To maximise scientific advance and support the Ontology's development and dissemination, we will convene an International BCI Ontology Consortium. Modelled on the influential CONSORT group, this will ensure cross-national engagement and cultural sensitivity. It will aggregate evidence (e.g. about reliability data, range of applications, need for adaptations across contexts), release updated versions as appropriate, and disseminate beyond countries currently using the method.

Deliverables:

(i) A published BCI Ontology including taxonomies of all key components

(ii) A training and testing set of research reports, annotated according to the BCI Ontology

(iii) An international consortium to engage and co-ordinate the scientific and applied communities in developing and updating the BCI Ontology.

WP2: From passive organisation to active interpretation of, and learning from, evidence

WP1 lays the groundwork for WP2, though the two will move forward in parallel, each informing the other.

WP2 involves creation of the AI System. One part of this will be the development and training of computerised NLP algorithms to extract key features of studies and use them to 'populate' the BCI Ontology with data, i.e. to fill the database with details of the studies. This populated Ontology will include the key components of BCIs, usage, putative mechanisms of action and context, plus evidence of behavioural outcomes and effects, together with study features that may affect the reliability of the findings.

The second component of the AI System will be a Reasoning System to draw inferences from the populated Ontology and create and revise appropriate knowledge structures. Machine Learning algorithms will continually revise and extend the knowledge structures and the Reasoning System as new information is fed in. For technical details, see Annexe.

.

The AI System will go beyond storing and retrieving information involving existing knowledge to making inferences about new interventions, settings, populations and behaviours. It will be a tool for both predicting and explaining BCI effects.

Thus WP2 comprises two main tasks involving the development of:

1. **Natural Language Processing algorithms to read the research literature**

These algorithms will be trained by BC experts to extract key features from research reports in a form that can be used to populate the BCI Ontology. The starting point will be the set of 300+ expert-annotated research reports from WP1. The accuracy of extraction of BCI Ontology features will be assessed against the BC expert-annotated test set of research reports. Study features that may influence the reliability of intervention effects will be annotated, including study design (using EQUATOR guidelines[20]), quality of outcome measures, missing data etc. These features will be used to assess risk of bias and hence confidence in reported effects[21,22].

2. **A Reasoning System and knowledge structures**

The populated BCI Ontology will provide input into the Reasoning System which will be developed to interpret the data in order to make generalisations, derive explanations and formulate predictions. The Reasoning System will be developed from similar systems that have proved successful in other areas[23,24], and will also use more broadly based knowledge gleaned from the BC literature (e.g. relating to the effects of reward and punishment on behaviour, factors influencing ability to exercise self-control, the development of habits, models of choice and decision making etc).

The Reasoning System will go far beyond storing and retrieving information about studies, and will make predictions in areas where currently little or no data exist. The system's predictions can then be compared with new data that are fed in to continually refine its knowledge structure and reasoning processes so as to minimise discrepancies. In addition to comparing the Reasoning System's predictions with new data, its results will be evaluated by BC experts and users for accuracy and intelligibility. This information will be used to continue training the Reasoning System to improve its performance, via the user interface (see WP3). Thus the Reasoning System will constantly evolve and improve in accuracy and scope.

In developing the Reasoning System, we will learn from similar exercises in other areas of study. For example, "4P Medicine" (Predictive, Preventive, Personalized, and Participatory) is fuelled by "systems approaches" to disease, emerging technologies and analytical tools.[8,9] 4P Medicine generates interpretations of personalised data permitting more diagnoses and treatment options to be considered than a medical practitioner might have thought of. Such systems are able to detect, derive and explain potential diagnoses, which are not discoverable by humans.

Deliverables:

(i) A set of NLP algorithms to extract and organise features from research reports in a form that can be used to populate the BCI Ontology.
(ii) A Reasoning System to create knowledge structures representing BC generalisations, explanations and predictions, to be refined and updated continually using Machine Learning as new evidence is fed in.

WP3: Using the BCI Ontology and AI System: real-time evidence accumulation, evaluation and dissemination via a user interface

WP3 will make the knowledge generated in WP1 and WP2 available (i) for practical application and decision-making in developing intervention policy and practice, and (ii) to further our understanding of the science of BC. WP3 will also evaluate the usefulness and validity of the BCI Ontology and AI System.

WP3 involves 5 tasks:

1. **Development of a user interface** to interrogate and provide feedback on the knowledge structures developed by the AI System, as well as providing for an intelligent search of the BC literature as characterised using the BCI Ontology. This will take the form of an online portal.

   Users will be able to specify details of their BC problems (e.g. population features, outcome domains, and intervention options) and be able to identify: (i) previous research on the efficacy of different intervention strategies; and (ii) suggestions from the Reasoning System in WP2 as to what intervention approaches may be useful, even if they have not yet been evaluated in the specific context or for that particular behavioural outcome.

   For example, a UK Local Government public health department might interrogate the system for the best evidence about interventions to reduce smoking initiation in their area with particular characteristics in terms of geography, social disadvantage, and adult smoking patterns. They would expect to learn what techniques should be included and how, when and where they should be delivered, how such an intervention might have its effect over what time period, and how big an effect might be expected. This would contrast with current practice where a limited set of already outdated evidence in contexts of unknown comparability is used to inform evidence-based recommendations. The online portal will enable the scientific outcomes of WP1 and WP2 to be made available for use globally, facilitating further convergence of the field around the new BCI Ontology, and demonstrating the ability of new computer science technologies to reproduce human reasoning at scale.

2. **User evaluation of online portal**: As well as a means of dissemination, the online portal will be a crucial means of testing the validity and utility of the Ontology and its application. We will use feedback mechanisms in the portal in order to conduct the evaluation by:
   a. undertaking formal user evaluation through observation, ‘think-aloud’[25], interview and focus group studies. While feedback on the user interface will be valuable, and inform its development, the main focus of this evaluative work will be to understand how users interpret the BCI Ontology and use the portal to develop intervention solutions for their own contexts.
   b. analysing user ‘click-through’ paths in the portal, to understand how the portal is used and identify important ‘end points’ – where users have either found the answers they were looking for, or have left with questions unanswered.
   c. seeking user feedback through prompts and surveys (accepting the self-selecting and unpredictable nature of these types of data, but using information here to generate questions for other parts of the evaluation).

3. **Pilot meta-analyses using the online portal**: We will pilot the ability of the AI System to maintain continuously updated synthesis of specific interventions[26], which include synthesis of statistical outcome data, and conduct meta-analyses on demand, based on

user-selections from the Ontology. (We will be cautious with regards to the practicability and desirability of such novel functionality.)

4. **Evaluation of BCI Ontology and AI System:** We will conduct an evaluation through the online portal and through case studies, selecting one topic where we expect the literature and ontological concepts to be relatively coherent (smoking); and a less easily characterised area, and thus more challenging to model successfully (alcohol, diet or activity). In each case study, in collaboration with BC experts in that specific area, we will undertake evidence synthesis and intervention development activities, assessing the strengths and weaknesses of the component parts of the evidence system on the following evaluation criteria.
    a. *The adequacy of the new system in comparison with traditional evidence synthesis*. We will: a) create automated systematic reviews using the Ontology to select relevant studies in conjunction with user input; b) use the automated data extraction and risk of bias assessment tools to conduct syntheses; and c) compare the results of this computer-assisted work with published systematic reviews, evaluating the automated reviews in terms of recall (are all the correct studies identified?), descriptive accuracy (are the studies correctly described, and risk of bias correctly assessed?), and inferential claims (how do the conclusions compare with those from manually-conducted systematic reviews?).
    b. *The performance of the feature extraction of the BCI Ontology against new evidence* (i.e. whether users agree with the assignments made automatically);
    c. *The salience and validity of the outputs of the AI System* as assessed by users;
    d. *The utility of the entire system* as assessed by decision-makers, initially for domains with simple behaviours, robust outcome measures and relatively coherent evidence.

5. **Implementation**: Effective implementation requires buy-in from collaborators from the start and we have already had positive responses from a range of users e.g. from public policy makers (Public Health England), commercial (Bupa, Nuffield Health), journal publishing (Addiction; Implementation Science; Annals of Behavioral Medicine) and funding (CRUK, MRC) sectors. We will work with funders and intervention repositories to ensure that the BCI Ontology becomes a *de facto* standard to report findings from BC research and applications. Enthusiasm for collaboration is evidenced by letters of support from the National Institute for Health and Care Excellence (NICE) and the US National Institutes of Health (NIH), offering supplemental funding.

Deliverables:

(i) An online, open-access, dynamic portal enabling
    a. scientists to generate new hypotheses and insights about BC; and
    b. intervention designers, practitioners and policy-makers to assemble optimised BC interventions

(ii) An ontology-based BC evidence query system
    a. verified in the domains of smoking, alcohol, diet and activity
    b. compared with traditional evidence synthesis on four criteria.

(iii) An evaluation of the system compared with traditional evidence synthesis on four criteria.

D. Timetable and milestones

The work flow is shown in the Chart and Table below.