

**D2.2: User Challenge 2**

**Early detection and characterisation of new or newly resistant/virulent pathogens using genomics and post genomics requiring laboratory-based facilities.**

Grand Challenge: to develop a laboratory-based system using genomic- and post-genomic methods to detect and characterise, rapidly and cost-effectively, all new and newly resistant virulent pathogens.

*This report has been commissioned as part of the UK Government's Foresight project, Infectious Diseases: preparing for the future. The views expressed do not represent the policy of any government or organisation.*

## **User Challenge 2 - Champion**

Dr Debby Reynolds, Chief Veterinary Officer, Defra

### **User Challenge 2 - Working Group members and contributors**

Dr Ian Barker (Technical lead)	Central Science Laboratory
Professor Joe Brownlie (Foresight Science Co-ordination Group)	Royal Veterinary College
Ann Davison	Defra consumer engagement project
Dr Donald King	Institute for Animal Health
Dr Stephen Little	DxS Ltd
Dr Anne O'Garra	National Institute for Medical Research
Dr Julian Parkhill	The Sanger Institute
Dr William Otim-Nape	National Agricultural Research Organisation, Uganda
Dr Mark Rweyemamu	Consultant, Tanzania
Dr David Serwadda	Makerere University, Institute of Public Health, Uganda
Dr Nicholas Thomson	The Sanger Institute
Dr Penny Wilson	Independent Consultant
Dr Simon Wood	Labformatics Ltd

### **Commissioned Reviews:**

1. The Implications of early detection of HIV (Professor Tony Barnett, LSE)
2. Ethical and Social Issues (Dr Richard Ashcroft, Imperial College London)

## Introduction

This report considers the opportunities and challenges offered by new DIM (Detection, Identification and Monitoring) technology from a UK perspective but includes a commentary on the situation in sub-Saharan Africa (SSA).

The potential for the emergence of new pathogens and the threat of pandemic disease and drug resistance underlines the need for effective disease surveillance systems. Predicted future advances in DIM technologies offer unparalleled opportunities to improve our ability to detect and characterise newly emerging pathogens in a timely manner. Early detection will in turn enable government and industry to plan and deploy more effective control strategies such as the development of specific vaccines. This challenge focuses on the characterisation of new and emerging pathogens rather than routine diagnostic laboratory functions. We envisage that the technology will be predominantly laboratory based although we recognise that laboratories could include both central, regional and even mobile models. In the longer term, some such capability may move out into the 'field' (see UC3; hand-held devices), but the need for laboratory facilities to do investigational work is likely to remain for the foreseeable future.

Two complementary approaches are envisaged. Host (immune) response signals in animals and humans will be interpreted to indicate likely causal agents and may also usefully give additional information as to the class of pathogen involved e.g. viral or bacterial etc. Host response may be a less useful tool in characterising plant disease. Pathogen genomics and bio-informatics will also open up new opportunities in characterising new and emerging diseases. Both approaches are likely to utilise sequencing and array (both nucleic acid and protein) based technologies. There is a strong commercial imperative to develop the necessary hardware platforms, but these are not likely to be interoperable and will represent competing technologies. They may not all be 'open-access' systems. Market size, volume and value of result will dictate whether commercial companies develop arrays and other reagents to aid the characterisation of infectious disease and the whole spectrum of interest (from the public standpoint) may not be catered for. It is clear that there are large gaps in our background knowledge in both immune signatures of disease, the composition of 'normal' microbial communities and the ability to predict pathogenicity and resistance characteristics from novel sequence data. Social issues surround concerns of informed consent and data privacy along with the need to raise the awareness of the importance of characterising new or emerging pathogens amongst the various interested parties. The capacity of sub-Saharan African (SSA) countries to operate such systems will also be a challenge but one which needs addressing.

## **1. Key future capabilities for user challenge system 2**

The aim of this section is to identify what capabilities for future DIM could become available in the future; when they could become available; and to relate these capabilities to the development of the underlying science and technology.

The scientific infrastructure behind UC2, that of identifying novel pathogens and variants, will fundamentally rely on continuing laboratory investigations examining and comparing the genomes of large numbers of viral, bacterial and fungal pathogens and non-pathogens. This work will be significantly enhanced by current and ongoing technological advances in sequencing and microarray capacity, primarily driven by the commercial imperative to rapidly and cheaply identify global and specific differences between individual human genomes. The ability to very rapidly generate de-novo and comparative genome information from a very wide range of pathogens and non-pathogens will become routine within five to ten years. This will enable the generation of the very large fundamental data sets that will be needed for the accurate and sensitive detection of novel pathogens and pathogenic variants using genomic sequence data. These datasets will need to be interoperable on a global scale and will require international co-operation, both to share data and to develop standard protocols for data storage.

The forward edge of UC2, the detection and identification of novel pathogens and variants in the field (20 years plus), will depend on the miniaturisation and automation of some of these hybridisation and sequencing technologies. It is likely that hybridisation (microarray) and PCR (polymerase chain reaction)-based technology will supplement or fully replace current technologies for identification and characterisation of known agents in the medium term (perhaps ten to 20 years). These technologies require information on the sequence of known agents, but are likely to be more readily produced in kit form for standardised machines, or in hand-held devices. Detection of unknowns can be enhanced by using a hierarchical approach; non-specific probes can be used to identify the organism at a very general level, with progressively more specific probes used to narrow down the identity. Greater use of sequencing, re-sequencing and microarray-based comparative genomic hybridisation in the research arena will massively increase knowledge of microbial variation and epidemiology. This will allow the continual refinement of detection and identification systems based on such prior knowledge.

Beyond this, novel sequencing technologies, based on very-high-throughput random sequencing or single-molecule sequencing (perhaps at the nano-scale), may well become the method of choice for identifying and classifying disease agents. These require no prior knowledge of individual sequences and should allow unusual or completely novel agents to be identified, irrespective of diagnosis. Incorporating this technology into clinical laboratories or field devices would be an enormous step forward for both medicine and research, and should be expected towards the end of the 25-year timescale of Foresight.

The continuation of large-scale microbial genome sequencing, particularly within-species comparative sequencing, has begun to identify some general evolutionary pathways followed by certain recently-evolved human pathogens. Current studies have identified plasmids and other mobile chromosomal elements that are associated with acquired drug resistance, adaptation to specific pathogenic niches and a heightened pathogenic potential. While it may be some time before this becomes predictive, it can be envisaged that such sequences could be monitored for their occurrence in novel chromosomal contexts. The increase in whole genome sequencing suggested above would allow much more than just individual diagnoses. It would be essential for detection devices to be connected to a centralised database, ideally in real time. This would allow up-to-date diagnosis, but, more importantly, collecting and monitoring the genome sequences of organisms in the field should facilitate the study of evolutionary processes as they occur.

The long term technical challenge is to develop the technical capability and tools to be able to predict the significance of novel pathogen sequences and variation in terms of altered virulence, host range, drug resistance etc. This challenge also includes the need to be able to distinguish significant genomic signatures from the background of the 'normal' plant, animal or human microbiota, particularly in non-sterile samples (such as stool samples). In the short to medium term it is likely that both array and random high-throughput sequencing based approaches will have to be used in a complementary fashion to achieve the dual goals of detecting both newly-emergent agents and detection of specific changes in organisms that cause clinical disease. In the very long term, massively parallel and low cost sequencing, coupled with powerful bioinformatics tools and global databases will slowly combine to render this distinction less important.

Infection of an animal host with a pathogen results in rapid changes in gene expression and protein synthesis by cells of the host's immune response. Collectively we refer to these changes as the 'immune signature of the infection'. The measurement of these immune signatures composed of gene expression profiles in blood leucocytes by microarray and/or real time PCR, together with analysis of specific immune mediators in serum, offers an attractive and complementary method for rapid detection and characterisation of unknown infectious pathogens. These immune signatures, once determined, will be able to identify the class of pathogen, for example, virus, bacteria or parasite, and may also be able to identify the specific microorganism, and even a mixture of organisms causative of the disease. This may be particularly valuable when the pathogen is unknown, or undetectable in readily-accessible host bodily fluids. In these situations the host signature may at the least be able to determine whether the infectious agent is a virus, bacteria or parasite, thus providing rapid knowledge for therapy and the type of containment advisable. Ideally microbial signatures would be determined at the same time as host immune signatures.

Microarrays for analyses of host gene expression most used to date are oligonucleotide arrays which have very good reproducibility, but because of the method of production and patent protection, the costs are prohibitive for

large-scale screening of immune signatures. This approach is now being overtaken by BeadChip microarray technology, and this is anticipated to dramatically reduce the cost of genome-wide expression analysis and allow researchers to expand the scale of biological experimentation, and increase the throughput. However, it will be at least five years before this technology is being used widely due to the cost for replacement of the equipment as well as comparison with known microarray approaches. Once the classifier genes have been identified, defining the specific signature resulting from infection with a particular pathogen, it will then be possible to run chips containing fewer transcripts and thus reduce costs further, and also reduce the complexity of assimilation of data dramatically. It is also possible the PCR-based technology will supplement this approach.

A major research undertaking is envisaged in the coming years to realise the vision of using immune signatures to diagnose infectious diseases. This will require acquisition of samples of blood from hosts infected with every known pathogen at different stages after infection and/or cure and their analysis by microarray and protein determination. It is envisaged that the acquisition of gene expression data by microarray will take a least ten years for the main infectious pathogens of humans. Microarrays for the genome of other animals such as cows, sheep and birds are clearly not as readily advanced as those for humans and thus it is anticipated that the compilation of such data will follow at a slower pace (possibly ten to 20 years).

Since different microorganisms induce distinct sets of soluble mediators, it is also possible to use the levels and signatures of these proteins in the serum, sputum or tears to identify at least the type of infectious pathogen. A rapid and high throughput method for simultaneously detecting multiple known soluble factors is suspension array technology is currently available but as yet is limited in scope even in human systems. The ultimate goal of a proteomic approach for analysis of host protein immune signatures analogous to the microarray approach is as yet limited by technical limitations in sensitivity and reproducibility and may only come into the mainstream in the next five to 20 years. The suspension array technology can also be used to measure microbe-specific antibodies in serum, sputum or tears. Effective detection and diagnosis of infectious disease, using serological techniques, usually requires testing for the presence of multiple antibodies directed against the pathogen and such multiplexed tests are already available for a large number of pathogens. These will need to be broadly expanded which will probably be effected within the next five years for known pathogens, but, as yet have no applicability with unknown pathogens. The attraction of these multiplexed tests is that they may deliver more economical and accurate results than is possible with traditional methods.

The compilation of databases containing such immune signatures is essential to enable diagnosis of the type of pathogen e.g. viruses, bacteria, parasites. This will also be able to distinguish different stages and types of immune responses to pathogens, e.g. ongoing infection, successful clearance, chronic infection, immune pathology. For the use of immune signatures to become a routine technique that can be applied globally in the field by non-experts, their

analysis will require the development of sophisticated bioinformatics and statistical methods that can handle large quantities of data, comparing them to known immune signatures in databases. Robust automatic and semi-automatic techniques will need to be developed to allow accurate diagnosis with minimal human intervention and easy exchange between centres of analyses. Although this technology is envisaged to advance rapidly (within the next five years) it is likely that the bottleneck will still be the acquisition of the data on immune signatures in infectious diseases of animals and man (five to 20 years).

## **2. The role of the systems in managing future risks**

The aim of this section is to place the future capabilities in the context of the future risks and the possible management of those risks.

It is inevitable that new pathogens affecting humans, livestock and plants will arise in the future. However, the relative contribution of different pathways for the emergence of these pathogens is unclear. In recent years there have been examples of novel diseases arising via a number of different sources and mechanisms. Principally these routes are: **(A)** either from previously unknown wildlife/animal reservoirs (such as HIV and SARS) or **(B)** phenotypic modification of a known agent through (i) genome mutation (influenza), (ii) uptake of resistance genes (antibiotic resistance in bacteria or fungicide resistance in plant pathogenic fungi), (iii) genome re-assortment (human-avian influenza) and (iv) linked and unlinked recombination (Western equine encephalitis virus and BVDV respectively). Fortunately, other potential avenues such as deliberate release of an engineered bio-weapon or uncontrolled escape of virulent laboratory-modified agents have not yet occurred, although it is possible that these routes of introduction could have increasing significance in the future. Established methods for identifying novel agents rely largely on prior knowledge of the antigenic or genetic characteristics of the pathogen (or related pathogens) and the use of models to understand the determinants responsible for pathogenicity. In the absence of a specific assay, highly conserved genome targets (such as 18s rRNA and 16 srRNA for eukaryotes and prokaryotes respectively) can be used to detect novel organisms in the environment. Analogous approaches based on assays that target common gene functions such as critical enzymes involved in replication may also be appropriate for viruses. However, a major lesson from the BSE/nvCJD outbreak is that emerging pathogens may not necessarily conform to our expectations and therefore we need to accommodate infectious agents, which possess totally novel natural histories into future diagnostic capabilities.

This technical paper - UC2 - addresses the development of technology platforms that will be able to detect and characterise new and emerging pathogens. We anticipate that this capability will complement (but probably not replace) routine diagnostic functions for the detection of individual known infectious agents. Such an ability to monitor the environment for the presence of novel pathogens (as well as new variants of known agents) will have a

number of benefits that will impact upon the epidemiology and strategies used to control infectious diseases. Firstly, we would envisage that the information generated from these assays would be fundamental for the implementation of effective measures to control and eradicate diseases; such as the selection of appropriate vaccines or the rational design of antimicrobials. Secondly, the ability to generate high-resolution characterisation data (such as full genome sequences for a rapidly evolving RNA virus such as FMDV) will advance our ability to monitor the transmission of disease and enable us to trace the routes of infection during an outbreak with high resolution. Furthermore, it is being increasingly recognised that many disease syndromes have complex aetiologies which result from the interplay between multiple infectious agents and which can also involve host factors and molecular triggers (e.g. PMWS). A system that would be able to multiplex both pathogen and host targets to present a 'bar-code' fingerprint would provide a paradigm-shift in our understanding of disease pathogenesis. 'Pathomics' (<http://www.llnl.gov/bio/groups/pathomics/>) is a term conceived to describe future detection schemes, which are based upon the mechanism of pathogenesis. It focuses on the changes in protein levels and other molecules that occur when a body has been exposed to a pathogen. Successful diagnostic strategies in the future are likely to utilise methods to detect and characterise both the infectious agent and the host response to infection.

### **3. The costs and benefits of selected future capabilities and their robustness to future uncertainty**

The aim of this section is to provide broad information, which will enable the UC champions and other high level stakeholders to appreciate the potential/relative importance of the future DIM capabilities.

A globally-connected system (consisting of shared pathogen/host genomic data in an interoperable format) would allow the construction of the very large database of baseline variation from which significant deviations could be detected, such as those driven by the introduction of a new therapeutic agent or the cross-species jump of a new or potential biological threat. A possible application where continuous and perhaps even fully-automated monitoring would be useful is in identifying new variants of the influenza virus. In this way, it might be possible to provide advanced warning of epidemics of novel or drug-resistant agents.

Large-scale integrated use of such systems would allow monitoring of pathogen populations, allowing the possibility of providing advanced warning of epidemics of novel or drug-resistant agents. The potential benefits of such an approach (e.g. early warning and more timely production of specific anti-viral vaccines) could be enormous. It has been estimated that an outbreak of pandemic flu could cost the US economy between \$71.3 billion and \$165 billion alone (Centers for Disease Control and Prevention). Similarly, the SARS outbreak in 2003 was reported to have cost the global economy between \$10 and \$30 billion (Fresh Minds). Outbreaks of BSE in cattle in

Europe and the USA and FMD in the UK have also caused billions of pounds of losses in recent years. Emergent and recently spreading plant diseases threaten food security (*cassava mosaic virus* and banana wilt in Africa), native ecosystems (sudden oak death), disrupt trade pathways (potato brown rot) and the potential to cause economic losses of perhaps tens or hundreds of millions of pounds (potential establishment of soybean rust in US; threat of potato ring-rot to UK).

Tony Barnett, in a review commissioned for Foresight, estimates the global cost of the HIV/AIDS epidemic to be in the region of \$20 billion. This estimate does not take into account other costs in the non-formal economy and of course the human and social costs. The review considers the potential for intervention at a number of key stages in the epidemic and concludes that although benefits from adopting DIM technology would have been great, there is doubt as to whether intervention would have been effective or possible, particularly in the early stages. The author concludes that powerful interests and other barriers would most likely have prevented successful intervention at a very early stage and that zoonotic transfer was inescapable. However intervention to prevent the assumed process of the virus becoming more virulent as it moved out of Africa and around the world seems more plausible, and would have yielded enormous benefits. Significant scientific uncertainty surrounds the question of whether viral pathogenicity would have been high in Africa without this process. There remains a significant opportunity to apply DIM technology to managing the problem of resistance to anti-viral drugs. A key conclusion then is that successful future intervention in infectious disease, through the development and deployment of DIM technology, is dependent on the necessary political will, capacity and infrastructure to intervene rapidly and effectively.

The costs of implementing such surveillance systems have not been estimated but would be significant. It would be important, given competition for limited resources, that surveillance for novel or emerging pathogens was not conducted at the expense of routine monitoring of known threats. Resources could be targeted perhaps in the context of investigating clusters of patients with unusual symptoms as well as more systematic sampling on a geographical basis or high risk situations e.g. pathogen populations under antimicrobial (or pesticide) selection pressure. It might also be sensible to prioritise resources for enhanced surveillance into diseases where changes in human behaviour or production systems are thought more likely. The opportunities offered by enhanced disease surveillance will also differ depending on the nature of the disease in question e.g. the difference between highly and less transmissible agents. Efforts can be streamlined by co-ordinating the different organisations and systems involved in collecting data, including for wildlife.

Appropriate human resources need to be available to conduct testing. Regarding vets (and plant health specialists), a balance has to be struck between using an expensive resource for mundane testing tasks (diagnosis may become quite routine for some diseases) and maintaining the availability of professionals across the country and their regular presence on farms to

check farm health plans. This may involve appropriate use of paraprofessionals.

Significant uncertainty remains in our capability to recognise completely new classes of infectious agents, unrelated to known agents.

#### **4. The factors influencing the development of the future capabilities**

The aim here is to consider the most important factors that could affect the development of the future capabilities, and so identify/inform any actions which stakeholders might wish to take following the project to maximise public good.

The subject of this section of UC2 encompasses a series of related activities as outlined below. Some of these will be mainly influenced by technological factors whereas for others epidemiological or commercial issues will be more important. The six steps implicit in the title of this section are:

- 1 Identification of the emergence of a novel or newly resistant pathogen
- 2 Isolation of the disease-causing organism
- 3 Analysis of the pathogen and its effects on the host
- 4 Selection of diagnostic markers to enable the specific and sensitive detection of the pathogen
- 5 Development of diagnostic tools to measure these markers
- 6 Deployment of diagnostic assays to detect the novel pathogen

The six-step process can be broadly split into three phases each with its own key factors influencing its development.

##### **Phase I – Identification**

It is prerequisite of any search for a pathogen that a novel disease or disease pattern has been recognised. Whereas for newly acquired drug resistance this is a fairly straightforward process, the realisation that a novel disease has emerged is a more complex issue. This public health role is not currently performed by commercial organisations and this situation is unlikely to change in the foreseeable future (Note: a case can be made for the private sector to share costs in the animal and plant health sectors). The key factors influencing development is continued support for the national and international organisations charged with the mandate for disease surveillance and the development of technologies that can identify that a new disease has emerged. The latter comprises methods for direct measurement of the infectious agent and methods for measurement of host response. There is the potential for some of these analysis tools to be commercialised but in the absence of a major predictable market they will likely appear as research tools as opposed to diagnostic products.

## **Phase II – Isolation and Analysis**

The desired time-scale for isolating and analysing novel pathogens is very short. Ideally once a new disease has been identified, the causative agent should be identified within weeks or even days. Given the sporadic nature of the requirement for this capability it is unlikely that a commercial organisation would choose to provide such a service. The involvement of the commercial sector will be in the provision of analysis technology (DNA, RNA protein analysis systems) into the public sector to facilitate rapid isolation and analysis – in recent examples such as HIV and SARS this is the model that operated.

At this stage IP (intellectual property) considerations are unlikely to be important because these activities are primarily involved in discovery research.

The key technological factors influencing future development are two-fold:

Tools for obtaining data

Tools for understanding what the data means

The development of analysis technology needed to obtain data will be driven by the commercial sector as it attempts to develop more advanced tools to analyse DNA, RNA, proteins and metabolites. These tools will be generally applicable within many research fields (diagnostics, pharmaceuticals, biotechnology, etc) and these multiple applications will justify the considerable investment needed to bring them to the market.

Tools for data analysis may also be driven by the commercial sector but the commercial imperative for understanding certain diseases may not always be compelling. Where the type of analysis aligns with other areas of biological research – for example how to predict protein structure from gene sequence – it is likely that this will become commercially available, but specific databases and models relating to unusual diseases will be less commercially attractive.

Given the above analysis it seems likely that primary analysis technology will emerge without additional assistance and a key factor influencing development will be the development of methods for the translation of analytical data into diagnostic information.

## **Phase III - Diagnostic Marker Selection, Development and Deployment**

Whereas in the previous phase the commercial sector was tangentially involved through the provision of technology, in this phase it is likely to be directly involved through the development of diagnostic products. The main considerations of diagnostic providers are commercial – is the market big enough and who will pay for the products.

The key factor influencing this phase is the transition of information discovered by the public sector in phase II to the private sector in phase III.

The diagnostic industry will continue to develop its own analysis systems but a diagnostic comprises two elements – an analysis platform and specific reagents for the target agent. The key factor for the development of future diagnostic systems is to manage this transition so that it is commercially viable for a diagnostic provider to develop assays. There are a number of models for this transition – one is to make all the information from phase II publicly available so any provider can develop an assay. Alternatively the phase II information can be patented and licensed to a single supplier to encourage the development of lower volume assays which may not justify many different providers.

The key factor is having a strategy for managing the transfer of information into the private sector, which maximises the likelihood of reliable tests becoming available in the shortest possible time.

## **5. The barriers and enablers to implementing the key capabilities**

The aim here is to focus on the most important issues, which could affect the practical implementation of the future capabilities (once they have been developed), and therefore the amount of public good subsequently realised.

### **Commercial and resource perspectives**

As outlined in the previous section it is important to separate process of discovering diagnostic markers from that of delivering diagnostic tests.

As technology advances so will the ability to rapidly discover new diagnostic markers. Diagnostic marker discovery is essentially a research activity – it can be performed in one or just a few highly specialised centres and can rapidly adopt new methods as they become available. These centres can be publicly funded and can operate independently of market forces. New technology will continue to develop and this will improve the ability to identify new potential diagnostics.

In contrast the delivery of diagnostic tests is essentially a commercial activity. The following analysis makes the assumption that diagnostics will continue to be provided by the diagnostics industry – this means that factors that tend to discourage the commercial availability of an assay will be a barrier, and factors that encourage it will be enablers.

With this perspective it is clear that market size is a key consideration. For common diseases in wealthy countries market forces will ensure that diagnostics emerge from the industry, whereas for uncommon diseases in poorer countries the commercial case will be less convincing.

In light of this the key barrier to be overcome is how to make rare diseases and/or diseases of poorer countries commercially attractive. There are a number of solutions:

1 Encourage the development of technology to make the development and distribution of diagnostic tests low cost, so that a commercial return can be made on a smaller investment (explore the use of 'hybrid' (joint conventional and philanthropic) capital)

2 Grant a manufacturer a period of exclusivity in a market to justify the investment needed to bring the product to market

3 Separate the assay from the platform; a typical diagnostic system comprises two elements – a specific set of reagents for the disease in question and a platform, which is used to perform the analysis. The platforms tend to be proprietary to a single manufacturer. If open platforms were available it would reduce the size of investment to bring a specific product to market and would mean that products with smaller market sizes could become commercially viable. A programme to encourage the development of open diagnostic platforms is one route to overcome the commercial barrier of small markets. There is a possible role for public and philanthropic bodies to engage with the commercial diagnostic sector to promote such outcomes.

Human health is a clear priority for the public and for government regulation. However, as previously stated, a case can be made for sharing the costs required for improvements in surveillance of infectious disease in animal and plant health. Government funds are appropriate where the market fails to deliver the public interest, such as protecting the environment and animal welfare issues. Concerns about the environment and conditions in which food animals are raised are expressed by consumers, but come below issues such as sugar, salt and fat in food and food poisoning and do not always translate into willingness to pay extra in the shops. Similarly, general expressions of enthusiasm for food produced in England, maintenance of the countryside and diversity of wildlife are hard to quantify in financial terms.

The appropriate balance is being explored between the taxpayer and industry regarding the costs of animal health and welfare, and includes insurance issues at European level. In this context, some ranking of public concern/benefits is important. The case for taxpayer contribution may be stronger where clear public wishes, which may have costs for farmers, are demonstrable, such as protection of wildlife. Cost sharing is complicated by the difficulties in quantifying the costs and benefits of sampling and monitoring. Several EU countries use production levies to help spread and manage disease risk, and to enable industry to take ownership.

In the case of farms, a sense of partnership between government and stakeholders envisaged by the UK Animal Health and Welfare Strategy can also be regarded as an enabler. Potential levers to monitoring farm animals are legislation, cross-compliance, incentives through farm assurance and farm health schemes, and information and persuasion. Compensation may have perverse effects on compliance. Cost sharing can be used to maximise the incentive to prevent and contain animal disease. In Denmark, compensation

payments depend upon farm risk management. Similar considerations arise in plant health discussions although the issues are very complex.

There is limited information on the public's priorities. A more sophisticated debate with the public, including through opinion research should enable appropriate allocation of resources to infectious disease prevention. Consulting the public and/or their representatives can also create a better mutual understanding about risk, taking account of people's values and helping avoid inappropriate scares which divert resources. Available research indicates that consumers want information communicated simply but accurately, in a non-patronising way and in their own language.

### **Ethical and Social issues**

The principal ethical issues relating to UC2 can be given as follows (Ashcroft):

- The challenges of responding to individual patient needs when a novel or newly resistant or virulent pathogen is identified.
- Operating a surveillance system, which may involve invasive methods of sampling (such as taking blood samples). Less invasive methods such as sputum, urine, or stool samples may involve problems of social acceptance. Sampling involves a burden on participants which may not involve significant compensatory benefits, and some possible stigma or sense of lack of local control or involvement.
- Translation of findings into practice. At a minimum this would be to benefit those populations most affected. Of more concern would be the application of this data to restrict movement [or to impose trade restrictions in the case of animal and plant health]. The major ethical challenge, however, would be to use this data effectively to a) develop new treatments and prevention methods rapidly and b) ensure that that such new treatments and prevention methods are rapidly made available to poor or vulnerable populations (who may be the 'signal' populations for the emergence of the new pathogens in the first place).
- Control and access to data and the governance of the surveillance programmes. This data is likely to be of major local, national, international and global importance, and of value to both public and private sectors. Mechanisms for widespread sharing of relevant data will be needed, consistent with both data protection principles and a suitable incentive system for commercial enterprise in developing new treatments and prevention methods.

### **Perceptions about use of data**

The consequences of identification and monitoring for risk management and risk communication need to be planned and agreed in advance with stakeholders. For example, worries about the potential for combining various sources of farm and personal data need to be addressed.

## **Compliance**

Volunteers have come forward for the human genome project but there is considerable sensitivity over e.g. Aids and gender of babies' testing. A recent survey found 71% sceptical about genetic testing, of whom 34% feared that genetic tests would be used for unacceptable purposes. Issues of human compliance include privacy, choice, insurance implications and handling untreatable disease.

## **6. Issues of equity**

The aim of this section is to comment on the applicability and implications of the selected future capabilities to the developed world (sub-Saharan Africa in particular).

Significant issues of equity surround whether any new treatment and prevention measures are made rapidly available to poor or vulnerable populations (who may be the signal' populations for the emergence of new pathogens in the first place (Ashcroft). The uptake of DIM technology in sub-Saharan Africa will also depend on the development of low-cost solutions and the development of capacity, perhaps at first in regional centres of excellence. In particular:

- The early recognition of a new disease or a resistant organism requires that there be a sound knowledge of disease surveillance baseline and a governance system for notification of unusual diseases. The early identification of BSE in the UK was probably in the context of a good baseline knowledge of scrapie. So the new syndrome could be identified as unusual scrapie-like pathology in cattle, a species not known to be affected by scrapie. There was also a culture and a governance requirement to report unusual disease incidents. By contrast, it took nearly two years in China for SARS to be identified and reported, partly because there was no public health requirement for local public health authorities to report to the central authority of the occurrences of disease incidents outside a pre-determined list of notifiable diseases (Report D4.3). HIV virus was not readily identified initially in the animal host and AIDS was initially misdiagnosed in several places in Africa as malaria, pneumonia etc, simply because it was occurring in a region with a poor baseline disease diagnostic and monitoring services and in an environment of many infectious diseases generally diagnosed on the basis of clinical presentation rather than laboratory detection and identification. The cassava mosaic epidemic in the Great Lakes region took time to come to the attention of government, and by this time the epidemic had spread over a large area in Uganda. Even after its detection, the identification could not be done in Africa. It had to be taken to the UK because of lack of skills and capability to deal with the problem in Uganda. While in the UK, detection was only made possible with reference to existing databases of similar viruses. These points emphasize the need for a system of governance, a conducive policy environment and regulatory framework, infrastructure and capacity that could accelerate detection, identification and reporting of new infectious diseases.

- The DIID Africa workshop (A4) and the Africa synthesis reports (A1) have highlighted the weaknesses in the surveillance of human, animal and plant infectious diseases in Africa. These reports have called for measures to strengthen the national capacities for disease surveillance, the formation of inter-institutional networks as national virtual centres and the support for regional centres of excellence in infectious diseases, through what has been termed a Pan African Vision for Infectious Diseases (A1). This report also recognises the need for a new initiative of regional centres of excellence in infectious diseases. Nevertheless it is important to stress that effective disease detection and surveillance must be rooted first and foremost in national systems for DIM. The recognition of the new depends on its detection as distinct from the known baseline. Therefore strengthening DIM capacities in Africa is a prerequisite.
- Ultimately effective diagnosis and surveillance is likely to be a combination of the application of UC-3 hand-held devices at the field or provincial laboratory level and UC-2 laboratory based genomics at the national and regional levels. African scientists will need ready access to international genomic databases and this could come via the SMART partnerships and via regional and global centres of excellence as advocated in the African synthesis report (A1) and this report.
- Of paramount importance is getting African participation in this initiative as early as possible to bring about ownership, buy-in and relevance. For instance databases of sequences of African plant, animal and human pathogens and immune signatures of African animals and humans will need to be developed, integrated into global systems and made available for diagnosis of African problems. This will avoid the development of devices, which could end up targeting diseases of the developed world, to the disadvantage of Africa.
- If the proposal for UK-Africa smart partnerships were to be realised in the next five to ten years, the capacity for surveillance, isolation and analysis of disease causing agents – old and new – would be greatly improved, as would the global management of the risks posed by new or evolving disease causing agents. Some of the analysis would be performed *in situ* in Africa and the vast amount in the UK/developed country institutions. This would also reduce the kind of mistrust that often arises when experts from developed countries descend on the scene of a new outbreak in Africa, collect samples to be analysed in their home countries, leaving behind little residual impact on either routine DIM or disease risk management.
- Global surveillance and development of capacity in Africa will require funding from outside Africa. The development capacity for DIM for infectious diseases (whether new or emerging), especially in Africa, will require public and philanthropic international funding. The case needs to be made that such investment will benefit the whole world, hopefully a simpler task as infectious disease becomes increasingly recognised as a global issue.
- Surveys on future disease risks (T3) and the analysis of culture and governance issues (D4 and A1) have highlighted the impact of culture and governance on the effectiveness of DIM systems, especially in Africa. These barriers are likely to be mitigated by the proposed systems for inter-institutional networking and UK-Africa smart partnerships for infectious diseases that are advocated in the Africa Report (A1)

## 7. Summary and suggested points of action for Foresight

1 Current commercial interest and investment will deliver the technology platforms needed to achieve the challenge within the envisaged timeframe.

It may additionally be possible for stakeholders to engage with the commercial sector in SMART partnerships to encourage the development of 'open access' systems and low-cost solutions.

2 The ability to interpret sequence and immune signature data associated with new and newly emerging pathogens will lag behind the capability to generate it.

Investment is required in the base science, both in building sequence and immune signatures databases and in the ability to predict risk from novel pathogen sequence and immune signatures.

There is an opportunity for Foresight to broker an international debate on the development and sharing of interoperable pathogen genomic databases.

3 It is clear that the benefits expected from being able to rapidly characterise newly emerging pathogens are dependent on both securing public support, and having the necessary structures and political will in place, to intervene in a timely and effective manner.

A study (gap analysis) to inventory the current UK capacity and infrastructure (and how it can be better interconnected with the international effort) dedicated to characterising newly emergent pathogens (and systems to communicate findings to policy makers) would be valuable.

There remains a need to engage with the public (opinion research) to determine their priorities.

Need to develop DIM capacity first and foremost in national systems as well as regional centres of excellence in sub-Saharan Africa (along with strengthening the infrastructure and regulatory capacity to intervene). African participation in this initiative as early as possible to bring about ownership, buy-in and relevance is of paramount importance.

4. It is important that resources are not redirected to characterising newly emergent pathogens at the expense of the routine monitoring and surveillance of known threats. Achieving an appropriate balance between the two tasks requires careful thought in the context of limited resources.

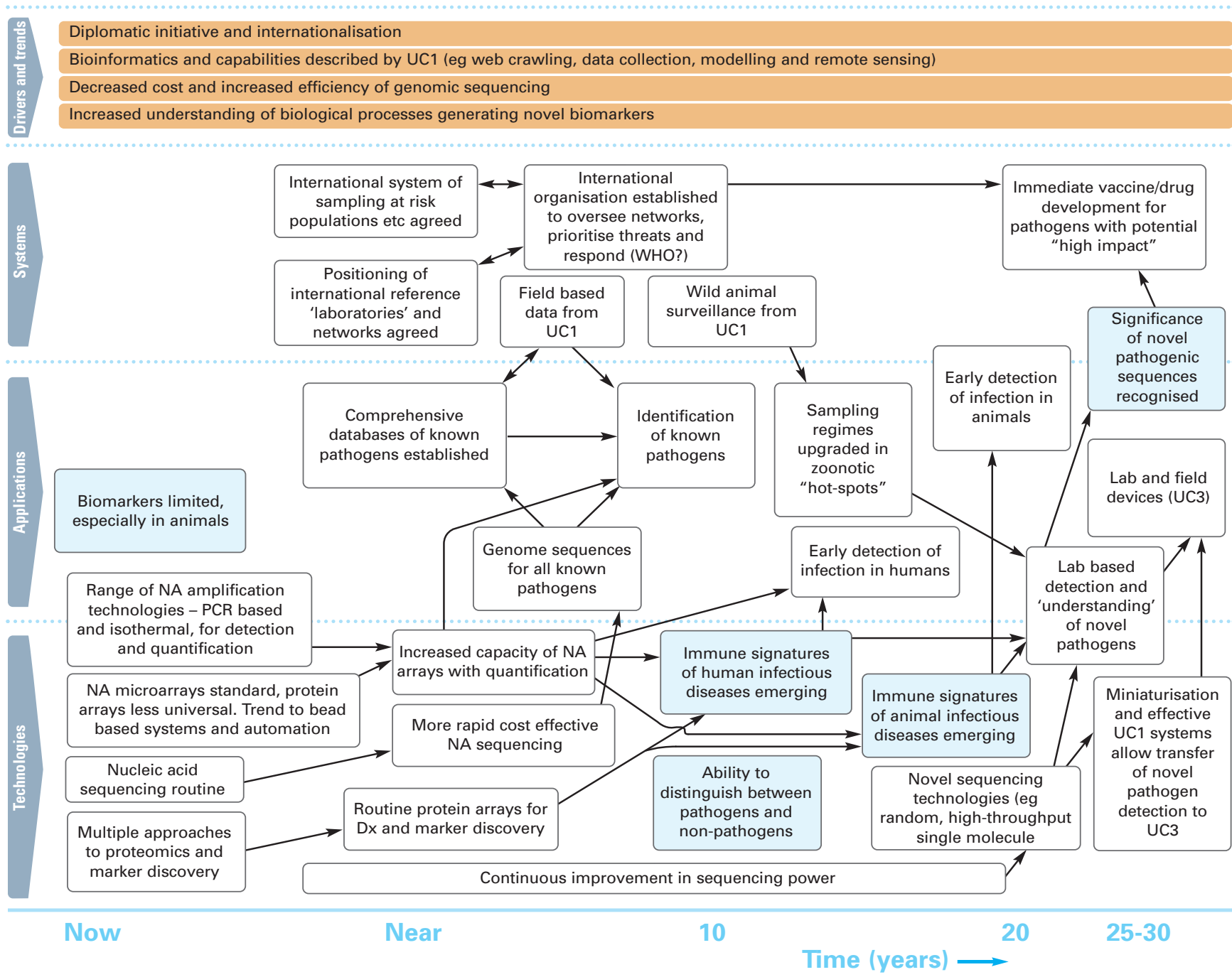
5. There are key issues of equity exist surrounding concepts of 'ownership of disease' (who pays/who benefits?), particularly in relation to animal and plant disease. An appropriate balance needs to be struck between industry and

government in supporting and resourcing pathogen surveillance. Numerous models including “shared cost” models can be envisaged.

6. The surveillance and monitoring of possible future bio-terrorism agents is best done as part of the same infrastructure which is used for ‘naturally occurring’ pathogens.

7. Mechanisms for widespread sharing of data will be needed, consistent with both data protection principles and a suitable incentive system for commercial enterprise in developing new treatments and prevention methods.

# UC2 Roadmap



Pale blue boxes represent capabilities that require a significant amount of fundamental research in order to be realised. The "Biomarkers limited, especially in animals" box has been highlighted as this is the starting point.

## **UC2 roadmap glossary**

(Terms that are not defined are deemed to be self-explanatory)

Pale-blue boxes in the roadmap represent capabilities that require a significant amount of fundamental research in order to be realised. The 'Biomarkers limited, especially in animals' box has been highlighted as this is the starting point.

### **Ability to distinguish between pathogens and non-pathogens**

This capability represents a significant UC2 milestone. In order to address this, it is necessary to understand what makes an organism pathogenic. Like many of the UC2 capabilities, a significant amount of fundamental research will be required to address this in full.

### **Bioinformatics and capabilities described by UC1 (e.g. web crawling, data collection, modelling and remote sensing)**

Bioinformatics and the global network capabilities described by UC1 will be imperative for the effective advance of UC2. While UC2 can initiate its own sampling regimes, information on potential hotspots from UC1 networks and systems will be critical. Management of the data generated by UC2 will rely on bioinformatics and UC1 databases.

Improvements in bioinformatics and the capabilities described by UC1 will occur independently of UC2 and help to drive UC2 forward.

### **Comprehensive databases of known pathogens established**

Database will be continuously improved and updated as research progresses. Existing databases should be combined and standardised.

### **Continuous improvement in sequencing power**

The efficiency of sequencing nucleic acids will continually improve – speed will increase and cost decrease.

### **Decreased cost and increased efficiency of genomic sequencing**

A capability likely to be funded and driven by research into areas other than infectious disease.

### **Diplomatic initiative and internationalisation**

Diplomatic initiative: the clearly necessary but very challenging drive to obtain widespread agreement in pursuit of a collective benefit.

Internationalisation: the drive to a connected world, in every sense. The global nature of infectious diseases means that their DIM and subsequent control can only be achieved through international agreements and networks.

### **Early detection of infection in humans**

The detection of pre-symptomatic disease may emerge from a comprehensive understanding of the immune signatures of disease.

In addition to detecting pre-symptomatic infections, it may be possible to ascertain an individual's/animal's/population's susceptibility to disease. Fundamental research is required to map single nucleotide polymorphisms (SNPs) and other genetic variations to susceptibility. Other factors, such as malnutrition, which is of particular significance in sub-Saharan Africa, will influence susceptibility and should form part of the bioinformatics algorithms etc.

### **Field-based data from UC1**

Data collected from web crawling and surveillance etc. begins to inform UC2.

### **Genome sequences for all known pathogens**

When one considers plants, animals and humans, this is an aggressive timeline. However, some data are already available and the predicted continuous improvement in sequencing power makes this possible.

### **Identification of known pathogens**

While known pathogens can be identified now, this box represents their rapid analysis and identification within a global reporting structure. At this point in the roadmap, routine sequencing and array technologies will be faster and less expensive. Bioinformatics and the comprehensive databases described above will play a significant role in achieving this goal.

### **Immediate vaccine/drug development for pathogens with potential high impact**

The early detection of novel pathogens will only be worthwhile if their identification results in an appropriate and measured response. This box represents the importance of immediate response to pathogens with potential high impact.

### **Immune signatures of animal infectious diseases emerging**

#### **Immune signatures of human infectious diseases emerging**

Infection of an animal or human host with a pathogen results in rapid changes in gene expression and protein synthesis by cells of the host's immune response. An appreciation of the immune signatures of disease will aid the detection of

pre-symptomatic infection and our understanding of pathogens and the disease process in general.

### **Increased capacity of nucleic acid arrays with quantification**

The use of nucleic acid microarrays are routine today. However, researchers make a choice to either look at the expression profiles in a qualitative manner, on microarrays, or measure the concentration of a number of defined nucleic acid sequences with, for example, quantitative polymerase chain reaction (PCR). Developments in analytical systems will bring these together. Initially, this capability is likely to be realised in a modular iterative system.

### **Increased understanding of biological processes generating novel biomarkers**

Fundamental research in areas outside infectious diseases, for example, in immunology and oncology, will provide novel candidate biomarkers to UC2.

### **International organisation established to oversee networks, prioritise threats and respond (WHO?)**

The international infrastructure required to oversee the implementation of the UC2 capabilities should be agreed at this point. This system may be an extension of the WHO's responsibilities. As UC2 matures, this and/or other organisation(s) will be paramount in prioritising threats and initiating appropriate responses, e.g. vaccine development.

### **International system of sampling at risk populations etc agreed**

The sampling regimes essential to UC2 will require international agreement. This may be challenging – it will be necessary, for example, to collect biological samples from wild animal populations.

### **Lab and field devices (UC3)**

It is generally accepted that, from a technological perspective, anything that is achievable in a laboratory setting ultimately has the potential, through miniaturisation etc, to be transferred to a hand-held device. Within the timeframe being considered by the Foresight project, it is unlikely that all the UC2 capabilities will be transferred to UC3 devices and systems. However, beyond 25 years, certain capabilities will become more field-based.

### **Lab-based detection and 'understanding' of novel pathogens**

Clearly, lab-based detection of pathogens is achievable today. However, attributing a pathogenic sequence to a particular disease can be time-consuming. This box represents the capability to identify novel pathogens and their significance in a timely manner. As portrayed in the roadmap, this

application will emerge from an understanding of the immune signatures of disease (at the nucleic acid and protein level), an ability to distinguish between pathogens and non-pathogens, and improved nucleic acid sequencing power. Comprehensive databases and bioinformatics systems will also play a critical role in reaching this milestone.

### **Miniaturisation and effective UC1 systems allow transfer of novel pathogen detection to UC3**

It is anticipated that the full realisation of this capability will extend beyond 25 years. The emphasis of the UC2 laboratory may change to include the co-ordination and regulation of DIM and the management of effective responses, e.g. vaccine development. Fundamental research into the biological processes of infectious disease will continue.

### **More rapid cost-effective nucleic acid sequencing**

This forms part of the continuous improvement in sequencing power. It has been boxed separately to define a timeframe.

### **Multiple approaches to proteomics and marker discovery**

This is the age of marker discovery. Today's 'bottleneck' is not in identifying candidate biomarkers but in validating them. Bioinformatics, proteomic approaches, nucleic acid microarrays and mass spectroscopy are all used routinely to discover novel biomarkers. In many cases, the relative concentrations of a series of markers will define a disease – this points to the continued need to establish excellent bioinformatics capabilities to analyse and manage complex algorithms.

### **Novel sequencing technologies (e.g. random, high-throughput single molecule)**

Continuous improvement in sequencing power will eventually result in a portfolio of rapid novel technologies. These will improve the efficiency of sequencing. Some may be suitable for UC3 devices, others will be more appropriate for high-throughput laboratory-based systems.

### **Nucleic acid microarrays standard, protein arrays less universal. Trend to bead-based systems and automation**

Microarrays are standard for the detection of DNA and RNA sequences. Protein arrays are technically more challenging but are in existence. As with many technologies, the trend is towards high-throughput screening and automation.

### **Nucleic acid sequencing routine**

A number of technologies and platforms are available for the routine sequencing of nucleic acids.

### **Positioning of international reference ‘laboratories’ and networks agreed**

The location of UC2 laboratories should be agreed alongside the sampling regimes. There are strong arguments for having a presence in Africa – perhaps a mobile laboratory moving between hotspots.

Discussions with leaders implementing UC1 capabilities should be initiated to ensure that the appropriate information is sourced, managed and delivered to UC2.

### **Range of nucleic acid amplification technologies – PCR-based and isothermal, for detection and quantification**

A wide range of nucleic acid amplification technologies are in routine use for the detection and quantification of nucleic acids. Techniques based on the polymerase chain reaction (PCR) are probably the most widespread. PCR methodology employs temperature cycling but a number of isothermal technologies exist.

### **Routine protein arrays for Dx and marker discovery**

It is anticipated that over the next 5 years some of the difficulties in making multiple protein measurements on single samples will be overcome. Protein arrays for diagnostics will become more routine. This advance in technology will also aid marker discovery.

### **Sampling regimes upgraded in zoonotic hotspots**

Wildlife surveillance from UC1 will inform UC2 of potential hotspots – sampling regimes can be introduced and/or upgraded in these areas.

### **Significance of novel pathogenic sequences recognised**

Beyond 25 years, it is anticipated that the full impact of a novel pathogen will be recognised from, for example, its sequence and immune signature. This knowledge will include an understanding of where mutations are most likely to occur and the consequence of these mutations – e.g. mutations may be silent, promote drug resistance or produce a form of the pathogen likely to make a species jump.

### **Wild animal surveillance from UC1**

Wildlife surveillance and data modelling from UC1 will inform UC2 of potential hotspots where pathogens are most likely – because of opportunity and environmental change etc. – to make the species jump from animal to human. In the first instance, UC2 may upgrade its screening regimes in hotspots. Through its databases and knowledge, UC2 experts will advise on the development of a diagnostic for the pathogen in humans, and on the effective rapid transfer of the test to UC3 devices.

All the reports and papers produced within the Foresight project 'Infectious Diseases: preparing for the future,' may be downloaded from the Foresight website ([www.foresight.gov.uk](http://www.foresight.gov.uk)). Requests for hard copies may also be made through this website.

First published April 2006. Department of Trade and Industry. [www.dti.gov.uk](http://www.dti.gov.uk)

© Crown copyright