Research Article

For reprint orders, please contact: reprints@futuremedicine.com

Journal of **Comparative Effectiveness Research**

Futurescapes: expectations in Europe for relative effectiveness evidence for drugs in 2020

Aim: Explore key factors influencing future expectations for the production of evidence of relative effectiveness (RE) for drugs in Europe in 2020; construct three plausible future scenarios for RE evidence generation. **Materials & methods:** Semi-structured key informant interviews and three rounds of modified Delphi to gather expert perspectives and develop future scenarios. **Results & conclusion:** Most influential factors were degree of regulator use of postmarketing authorization (postlaunch) efficacy studies and adaptive licensing; degree of pan-European health technology assessment body coordination in reviewing prelaunch evidence and demanding postlaunch studies; the nature of regulator – health technology assessment body interaction. The most likely scenario entailed some change with postlaunch regulatory studies driving the likely nature of RE evidence generated.

Keywords: adaptive licensing • coverage with evidence development • drug development • electronic health records • HTA methods • performance-based risk sharing agreements • relative effectiveness

Adrian Towse*[,1], Martina Garau[1], Penny Mohr[2] & Donna A Messner[3]

[1]Office of Health Economics, 7th Floor, Southside, 105 Victoria Street, Westminster, London, SW1E 6QT, UK

[2]Patient-Centered Outcomes Research Institute (PCORI), 1828 L Street, NW, 9th Floor, Washington, DC 20036, USA

[3]Center for Medical Technology Policy, World Trade Center Baltimore, 401 East Pratt Street, Suite 631, Baltimore, MD 21202, USA

*Author for correspondence:
Tel: +44 0 20 7747 1407
Fax: +44 0 20 7747 8851
atowse@ohe.org

Payer interest in relative effectiveness evidence

Patients and those paying for healthcare in the EU (termed Europe for the remainder of the paper) are increasingly interested in how much additional benefit or reduced side effects, one treatment offers as compared with another in routine clinical use. Two elements are of importance:

'Relative' or 'comparative' benefit. It is not enough that the therapy works. The question is "when does it add benefit as compared with existing treatments?"

'Effectiveness' rather than 'efficacy.' What matters to patients and payers is that this added benefit can be achieved in routine clinical use of the product as well as in controlled experimental conditions.

The European Commission's (EC) High Level Pharmaceutical Forum set out in 2008 [1] the need for greater use of relative effectiveness (RE) evidence in Europe to identify the value of pharmaceuticals to national healthcare systems. A series of programs funded by the EC and led by the European Network for health technology assessment (EUnetHTA) [2] have explored the potential for evidence to be generated and assessed in a consistent way across Europe by health technology assessment (HTA) bodies to reduce duplication of effort and achieve more rapid patient access to effective drugs. These HTA bodies act on behalf of national and regional payers in Europe to assess the value of drugs.

Regulator interest in RE evidence

Interest in RE has also been stimulated by the EMA, which assesses applications for marketing authorizations for drugs in Europe. It takes a life cycle approach to benefit risk assessment emphasizing postmarketing authorization (postlaunch) safety assessment of effectiveness [3]. This has led to development of the data infrastructure in Europe to monitor safety signals and to new pharmacovigilance legislation. This allows the EMA to require postauthorization efficacy studies (PAESs) [4]. Types of PAESs

Future Medicine part of fsg

that might be required [5] include: "Studies aimed at determining clinical outcome following initial assessment based on surrogate end points; studies on combinations with other medicinal products; studies in subpopulations; studies in the context of the European standard of care; studies aimed at determining long-term efficacy of a medicinal product; and studies in every day medical practice." The EMA is developing a Scientific Guideline to inform decisions as to whether or not to require a PAES and, if so, the type of study. When should it be an interventional study using randomized controlled trials (RCTs) or an observational or pragmatic clinical study?

Power to require PAESs (as well as postauthorization safety studies [PASSs]) may increase the EMA willingness to give earlier regulatory approval subject to additional data collection – conditional market authorization (CMA). This is currently restricted to particular patient groups and diseases and is often regarded by companies as a last resort if a full approval is unlikely to be obtained [6]. It may also lead to use of adaptive licensing (AL) – a structured process of review of the terms of drug approval in response to new evidence [7,8] – to provide earlier access for some patient populations to new drugs.

Regulator & payer interaction on RE evidence

Earlier regulatory approval will only lead to earlier patient access if national HTA bodies, acting on behalf of Europe's payers, are willing to reimburse drugs, at least conditionally, while more evidence is collected [9]. To date, the emphasis has been to improve the quality of the evidence in reimbursement submissions immediately following marketing authorization (at launch) through a EUnetHTA/EMA collaboration. This is to provide parallel scientific advice to companies on the design of clinical trials conducted before marketing authorization (prelaunch) to try and improve the quality of evidence being supplied at launch.

If the EMA discussed PAES study requirements with HTA bodies this may increase their willingness to conditionally approve reimbursement when the EMA gave conditional marketing authorization to a new drug. Such conditional approvals are often coverage with evidence development (CED) arrangements providing patient access while additional studies are completed, at which time another review can take place [10].

Drug development & RE evidence

This raises the question as to when RE evidence can be collected? Assessing effectiveness in clinical practice requires routine clinical use. However, some components of efficacy and effectiveness differences can be addressed in prelaunch pragmatic trials. The Innovative Medicines Initiative (IMI) of the EU and the European Federation of Pharmaceutical Industry Associations (EFPIA) is funding the 'get real' project [11] examining which 'real-world' data including RE evidence can be collected pre- and post-launch and by what study design.

These various initiatives may add to the cost of drug development, increasing prelaunch study complexity and adding postlaunch study requirements from regulators and HTA bodies. Patients may gain earlier access to new treatments and companies' earlier revenues. This could increase the cost pressures European payers' face. Payer interest in RE evidence reflects concern that treatments may not, in practice, provide good value. Yet if reimbursement is not achieved, manufacturers may not be willing to collect additional evidence. Drug development needs to respond to changing evidence requirements in a way that reduces development costs, or, at a minimum, stops them continuing to rise. In such an environment, drug use can be targeted at patients where treatment offers payers the greatest value, while still enabling companies to get a return on their research investment [12–14].

This paper reports the results of a future scenarios-building exercise for Europe. Alternative scenarios are designed to provide insights for drug development planning under differing expectations for RE evidence, pre- and post-launch, that will plausibly exist in 2020. Yet drug development is global and will not change solely because of developments in Europe. The US equivalent of RE – comparative effectiveness research (CER) – has gained momentum and scenarios for changing evidence requirements in the USA are set out in another paper [15] in this issue of the journal. A further paper in this issue [16] sets out similarities and differences in the future evidence expectations as between the EU and the USA. A subsequent paper develops a new drug development paradigm to address future EU and US environments.

Materials & methods

We created three distinct but plausible explorative scenarios of the possible future demand for, and ability to produce, RE evidence in Europe by 2020: a baseline or status quo scenario (one in which present conditions persist) with two alternative scenarios with different implications for the generation of RE evidence for drugs. The process used is described in the US scenarios paper [15]. It is briefly set out below.

We searched MEDLINE (PubMed) and gray literature (EC, EUnetHTA, EMA and others) for papers

on plans to develop RE evidence, and recommendations or predictions of change in the provision of RE evidence. We developed a preliminary list of prospective key factors (defined as "discrete events, developments, or trends which can be seen as having a significant positive or negative impact") depending on their future direction on: the demand for RE evidence; the technical feasibility to produce it; or the incentives to produce it for drugs in development.

We interviewed 13 key informants from regulatory bodies, HTA bodies, payers and academia, with expertise in public policy, regulatory science, HTA and clinical research methods. Each interview assessed the key factors, and explored alternative pathways along which those factors might develop. The first half of each interview was identical, the second portion was customized for the individual informant's expertise. We then refined and revised our list of key factors. Twelve key factors were chosen for evaluation in our modified Delphi approach.

This modified Delphi approach had the following stages:

- The first Delphi questionnaire was presented to an expert panel in an online format using Qualtrics© survey tools. Panelists ranked the overall importance of the 12 key factors. Additionally, they ranked the probability and impact of a set of potential future alternative pathways for each key factor;
- The most impactful and probable of these factor-specific alternative pathways (as scored on five-point Likert scales by the participants) were then built into four draft scenarios;
- The four scenarios were assessed during an in-person meeting (round 2 of the Delphi) with real-time voting (after clarification of the factors, pathways and scenarios) on both the probability of each of the four scenarios taking place and on the internal consistency of the elements within each scenario;
- Following the voting, two breakout groups with a balanced mix of experts refined the scenarios to be more internally consistent, and identifying missing elements. The full meeting discussed the results;
- Following the meeting, we constructed three final future scenarios. A final online vote was taken for each of the three scenarios (round 3 of the Delphi). The panel assessed the likelihood of each of three scenarios occurring by 2020 using a five-point Likert scale. Participants of the Delphi first and second rounds were invited to vote and a total of 13 experts voted.

The areas of expertise, of the key informants and modified Delphi participants at stages 1 and 2, are set out in Table 1. It was not possible to identify the composition of the respondents to round 3 of the Delphi. The pharmaceutical industry sponsors did not want to influence the initial formative parts of the process and were neither key informants nor in round 1. They did participate in the round 2 voting and discussion but not in the final round 3 vote.

Results

The 12 key factors generated from the key informant interviews, which we used in the first round of the modified Delphi panel survey are set out in Table 2.

When the participants in the first round of the modified Delphi exercise were asked to rank the likely importance of these factors, the results were as set out in Figure 1.

In response, we reduced the number of factors for use in constructing the scenarios to eight, grouping the eight factors into four themes:

- Regulatory. This combined the adaptive licensing and pharmacovigilance factors (fifth and seventh, respectively);
- HTA bodies/payers. This combined HTA demand for postlaunch studies (ranked first), assessment of clinical evidence (ranked third) and assessment of costs and decision-making (ranked fourth). In most systems, these are part of the same process;

Table 1. Informant and Delphi participants.

Expertise	Participants (n)		
	Key informants	Ranking (round 1)	Meeting (round 2)
Regulators	3	1	1
Academic and independent experts	3	8	5
Health technology assessment bodies and payers	7	3	5
Life sciences industry	–	–	6
Total	13	12	17

Table 2. Key factors for modified-Delphi round 1.

Key factors	Definition
New pharmacovigilance regulation	Pharmacovigilance rules are set up to prevent, detect and assess adverse reactions to medicines placed in the European market. The new regulation enables the EMA to seek 'efficacy' data in addition to 'safety' data in order to inform a benefit–risk assessment of a medicine in a postauthorization efficacy study
Adaptive licensing	Adaptive licensing is a prospectively planned, flexible approach to regulation of drugs and biologics. Through iterative phases of evidence gathering to reduce uncertainties followed by regulatory evaluation and license adaptation, adaptive licensing seeks to maximize the positive impact of new drugs on public health by balancing timely access for patients with the need to assess and to provide adequate evolving information on benefits and harms so that better – informed patient care decisions can be made [4]
Assessment of clinical evidence by HTA bodies	Scientific assessment of the incremental effectiveness of a new medicine (compared with current practice) based on clinical trial data and the use of modeling
Assessment of costs by HTA bodies or payers. Decision-making on price and/or reimbursement by HTA bodies or payers	Evaluation of cost or resource utilization implications of using a new medicine. Development of a decision on the appropriate use of a new medicine within the healthcare system or its price or reimbursement level
Demand for postlaunch RE studies by HTA bodies or payers	Studies requested by HTA bodies or payers to demonstrate benefits in real-world settings
Coordination between regulatory and HTA bodies	Interaction between regulatory and HTA bodies covering the offering of scientific advice to drug manufacturers. This could be pre- or post-launch, or both. It could lead to the coordination of, or agreement on, evidence requirements (e.g., the type of study design and type of end points to be included)
Use of PROs	The extent to which PROs will play a role in drug approval for licensing and for HTA review and payer reimbursement purposes. The choice of PROs could be either a generic or disease-specific instrument
Methods to analyze RE evidence	The extent to which robust methods to analyze evidence produced by RE studies will be developed and agreed by key stakeholders (e.g., regulatory bodies, HTA bodies, payers and the pharmaceutical industry). Areas of development include: – Methods for quantitative benefit-risk assessment – Methods to analyze data from observational studies – Methods to analyze data from pragmatic clinical trials – Methods to design and analyze adaptive trials – Analytical methods for data mining and analysis across multiple data sources enabling use of 'big data'
Infrastructures to conduct RE research	The extent to which improvements in research infrastructures and in the availability of data are made

HTA: Health technology assessment; PRO: Patient-reported outcome; RE: Relative effectiveness.

Table 2. Key factors for modified-Delphi round 1 (cont.).

Key factors	Definition
Relationship between the US FDA and EMA	Any interactions between FDA and EMA and possible alignments in their requirements on the following: – Willingness to use adaptive licensing – Conditions under which adaptive licensing is likely to be acceptable – Type of (postlaunch) studies required for benefit-risk assessment – Likelihood of requiring postlaunch efficacy/effectiveness studies (as opposed to a safety study) – Type of scientific advice given – Willingness to accept new methods such as adaptive trial design – Willingness to accept use of PROs, and the type of PRO required
Personalized medicine	The extent to which discoveries in genetics and genomics leads to drug development and therapeutics targeted to specific disease subtypes and patient groups, or to effective strategies for patient stratification
Commissioning and funding RE studies	Identification of which stakeholders will commission and pay for RE studies

HTA: Health technology assessment; PRO: Patient-reported outcome; RE: Relative effectiveness.

- Coordination between regulatory and HTA (ranked second);
- Infrastructures and methods. This combined methods to analyze RE data and RE research infrastructures, which came eighth and ninth, respectively.

We did not include the remaining four factors (use of patient-reported outcomes [PROs], personalized medicine, US FDA/EMA collaboration and commissioning and funding RE studies) in the preliminary scenarios. We set out in Table 3 below, the findings for the pathways within the eight factors in these four themes. The table shows the key informants assessment of the probability of occurrence of each pathway and the level of impact were that pathway to occur. We show figures for the aggregate of Likert scores 3–5, in other words, even odds to highly probable probability of occurrence, and moderate-to-high level of impact. We also show the mean scores on the 1–5 scale. Above 3.0 means above even odds or above moderate influence. The full results for each of the eight factors are set out in Supplementary Tables 1–8 (see online at www.futuremedicine.com/doi/suppl/10.2217/cer.15.7).

Four scenarios were then constructed to combine pathways from the eight key factors in alternative internally consistent ways of thinking about the future environment. These were voted on and discussed in the modified Delphi process round 2. In response to discussion in round 2, the factor 'commissioning and funding RE studies' (ranked tenth) was added to the infrastructure and methods theme.

Key factors

We therefore used nine factors to create the final three scenarios, grouped under the four main themes.

Regulatory

The introduction of PAESs. In the first scenario they are rarely used, leading to little change to the current arrangements. In scenarios 2 and 3, they are used regularly, potentially displacing any postlaunch requirements from HTA bodies. Companies focus on meeting PAES requirements and are reluctant to sign up to undertaking additional studies for HTA bodies without some element of agreement between HTA bodies, the EMA and the company as to the number and type of studies required. No such agreement occurs in scenario 2, but is possible in scenario 3.

The introduction of AL. In the first scenario, CMA is little used, as now, and seen by companies as inferior to getting an 'unconditional' market authorization. In scenario 2, the use of PAESs makes CMA more attractive to companies. If postlaunch study requirements

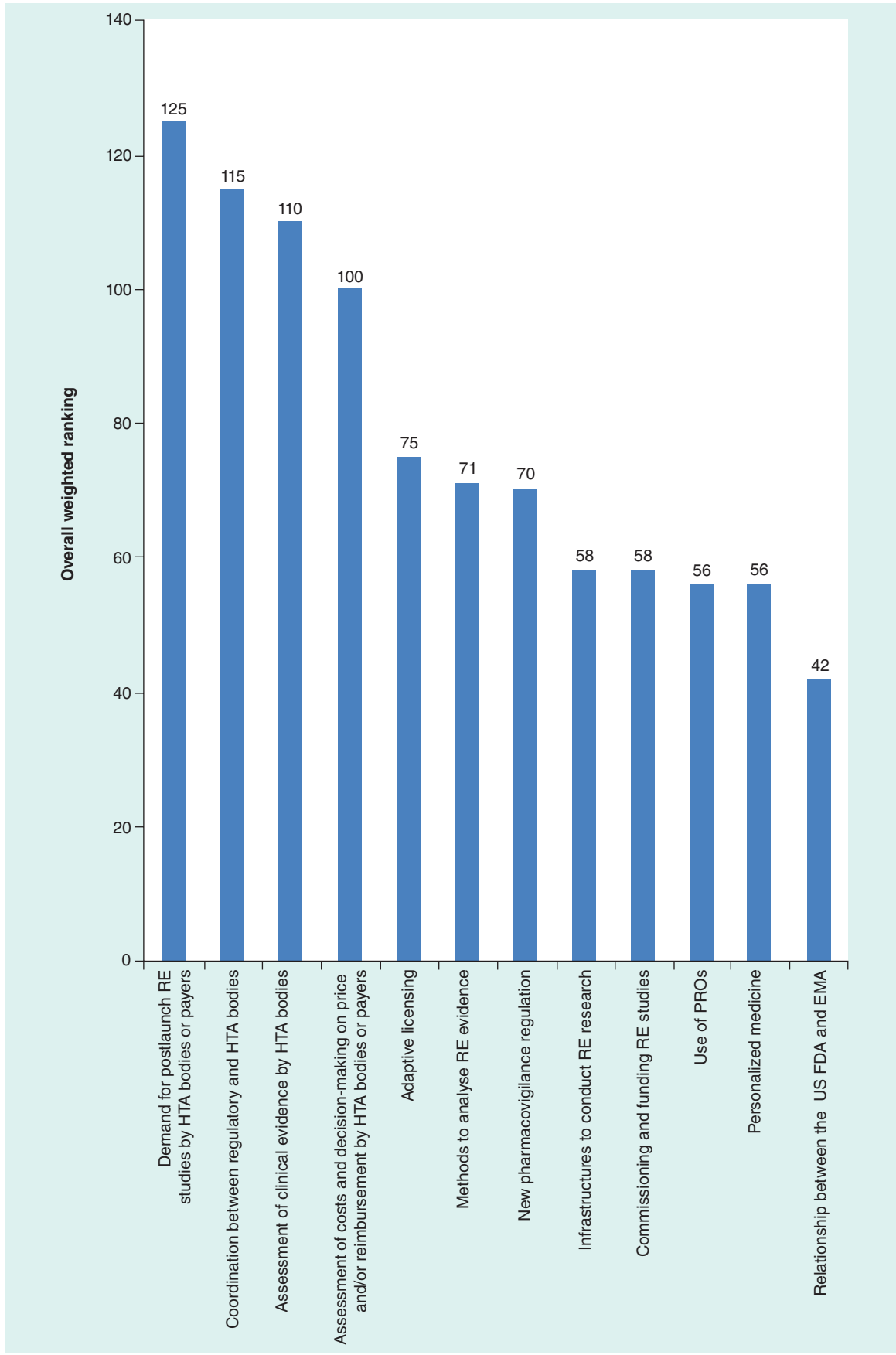
Overall weighted ranking
140
120
100
80
60
40
20
0
125
115
110
100
75
71
70
58
58
56
56
42
Demand for postlaunch RE studies by HTA bodies or payers
Coordination between regulatory and HTA bodies
Assessment of clinical evidence by HTA bodies
Assessment of costs and decision-making on price and/or reimbursement by HTA bodies or payers
Adaptive licensing
Methods to analyse RE evidence
New pharmacovigilance regulation
Infrastructures to conduct RE research
Commissioning and funding RE studies
Use of PROs
Personalized medicine
Relationship between the US FDA and EMA

Figure 1. Ranking of key factors (see facing page).
HTA: Health technology assessment; PRO: Patient-reported outcome; RE: Relative effectiveness.

are likely to be imposed by the EMA in any case, then achieving earlier access via a CMA makes sense. However, only in scenario 3 does the EMA directly link the use of PAESs to AL, in the sense of using postlaunch study requirements to explicitly address uncertainties at launch brought about by earlier licensing for some groups of patients. Here AL goes beyond the specifics of CMA, albeit without any requirement for new legislation because of its use of PAESs.

HTA bodies/payers

Methods for the assessment of RE evidence at launch. In scenario 1, while there is some information sharing across HTA bodies, there is no agreement on methods. EUnetHTA has little medium-term impact. There is no pan-European dossier submission based on the EUnetHTA Core Model® for rapid RE assessments (REAs) [17] and no shared assessment of the clinical data. In the second and third scenarios, EUnetHTA does impact. There is a European dossier and there is a shared process of review of the clinical evidence – in effect a form of mutual recognition along the lines of the processes proposed for the rapid REA pilots. Not all of the major member states participate however.

Location of decision-making or appraisal, and assessment of resource use, budget impact and affordability. These are assumed to remain the responsibility of the member state, and so be determined at national/regional level, in all three scenarios.

Payer/HTA body demand for postlaunch studies. Scenario 1 assumes that some countries continued to request postlaunch studies but there is no increase in demand. There is no cross-border coordination or willingness to accept observational or pragmatic clinical trial (PCT) data from studies conducted in other countries. In scenario 2, the advent of PAESs specified by the regulator makes companies less willing to agree to conduct separate postlaunch studies for HTA bodies. In scenario 3, however, there is an increase in demands for postlaunch studies from HTA bodies, with three additional elements: payers are more willing to enter into conditional reimbursement schemes, in other words, study requirements are linked to earlier access for patients (and earlier revenues for companies); there is dialogue between HTA bodies (and with the EMA) about the type of end points and studies that might be required; and there is some acceptance of study results from other countries.

Dialogue between regulators & HTA bodies/payers

Degree of coordination between the EMA and HTA bodies. In scenario 1, parallel scientific advice does not lead to any alignment of requirements. In scenario 2, there is some alignment of requirements for prelaunch studies. In scenario 3, coordination extends to postlaunch studies both in the sense that there is an attempt to align PAES requirements with HTA body requirements (in effect a postlaunch scientific dialogue) and the prelaunch dialogue addresses possible trade-offs between pre- and post-launch evidence collection.

Infrastructures, methods & funding for the conduct of RE studies

Improvements in infrastructure for the conduct of RE studies. In scenarios 1 and 2, progress is made in some countries in implementing electronic health records (EHRs), disease registries are available in some disease areas and other observational databases are developed. In scenario 3, there is more progress in implementing EHRs, which can facilitate the recruitment of, and monitoring of, patients in RCTs and PCTs, as well as enabling more observational studies to be conducted. There is more crossborder collaboration across large registries. Progress is made with 'big data' which presents new sources of data for generating evidence on RE, albeit ones that add to the challenges of demonstrating causality.

Progress on methods. In scenarios 1 and 2, there is limited progress on methods, which means that HTA bodies and regulators continue to have concerns about the robustness of some study designs, notably in the ability of observational studies to overcome problems of confounding. In scenario 3, more progress is made in addressing these concerns, which together with the availability of better quality data, leads to increased confidence in the results from such studies. All three scenarios assume that RCT efficacy evidence remains a core requirement of regulator and payers, with PCT and observational study evidence building on this.

Role of partnerships. In scenarios 1 and 2, industry is regarded as responsible for financing and organizing RE research. In scenario 3, a degree of collaboration occurs in designing, organizing and financing studies. Unlike the USA, there is no scenario in which payers organize the collection and analysis of RE data themselves.

The three final scenarios

Scenario 1 – no change; low trust environment

Scenario 1 is the 'base case' and a continuation of the status quo. We have characterized this also as being relatively low trust as between manufacturers, regulators and HTA bodies in their interactions with each other, and as between HTA bodies and their counterparts in

Table 3. Round 1 modified Delphi results for key factors.

Factors	Pathway number	Pathway description	Probability of occurrence (n = 13; scores 3–5)†	Occurrence mean score†	Level of Impact (n = 13; scores 3–5)‡	Level of impact mean score‡
New pharmaco-vigilance regulation	1	New legislation leads to little change in EMA practice. PASSs and PAESs may increase but they are kept separate. No specific EMA requirements for RE evidence postlaunch EU Commission proposal for postauthorization efficacy studies is finalized and integrated into implementation of the pharmacovigilance legislation	68%	3.31	22%	2.08
	2		75%	3.15	77%	3.85
Adaptive licensing	1	AL is only implemented and applied to a subset of drugs (e.g., oncology, orphan drugs or more generally where there is an unmet need). For certain therapeutic areas, companies still opt for a conventional regulatory approval route. There is a better use of current regulatory tools (e.g., conditional marketing approval) without the need for new legislation	91%	3.62	85%	3.15
	2	AL is implemented and applied to a broad range of drugs depending on their characteristics in order to enable use while additional studies are conducted to address one or more of the following: reduce uncertainty around an end point; broaden treatment-eligible population; reduce statistical uncertainty; enable new combination development; reduce uncertainty around study design; ensure 'real-world' effectiveness; address rare adverse events	37%	2.46	83%	3.69
	3	Implementation of AL proceeds slowly with no major change by 2020. Barriers such as EU Commission opposition and lack of manufacturer's commitment to do postlaunch studies are not overcome	69%	3.00	39%	1.92
Assessment of clinical evidence by HTA bodies and payers at launch	1	Although there is some collaboration and information sharing between HTA bodies, there is no formal agreement on methods (e.g., some countries accept indirect comparisons, others do not). EUnetHTA method guidelines are accepted only in certain countries (e.g., those which have not developed their own guidelines). Separate assessment of clinical evidence at national level	76%	3.23	53%	2.69

†1 = Very improbable; 2 = Improbable; 3 = Even odds; 4 = Probable; 5 = Highly probable.
‡1 = Slight-to-nil influence; 2 = Limited influence; 3 = Moderate influence; 4 = Considerable influence; 5 = Dominant influence.
AL: Adaptive licensing; EHR: Electronic health record; EUnetHTA: European Network for health technology assessment; HTA: Health technology assessment; PAES: Postauthorization efficacy study; PASS: Postauthorization safety study; PCT: Pragmatic clinical trial; RE: Relative effectiveness.

Table 3. Round 1 modified Delphi results for key factors (cont.).

Factors	Pathway number	Pathway description	Probability of occurrence (n = 13; scores 3–5)†	Occurrence mean score†	Level of Impact (n = 13; scores 3–5)‡	Level of impact mean score‡
Assessment of clinical evidence by HTA bodies and payers at launch (cont.)	2	There is a convergence of methods facilitated by EUnetHTA (e.g., on the appropriate use of network meta-analyses) and production of a pan-EU dossier (e.g., using efficacy data to conduct indirect comparisons). However, the pan-EU dossier is only accepted in certain countries. Others continue with separate national assessment	77%	3.38	67%	3.08
	3	There is a convergence of methods (facilitated by EUnetHTA) and production of a binding pan-EU dossier (also facilitated by EUnetHTA), in other words, no separate assessment of clinical evidence at national level	23%	1.77	53%	2.92
Assessment of costs and decision-making on pricing and reimbursement by HTA bodies and payers at launch	1	Separate assessment of cost (and, where required, cost–effectiveness) and separate decisions on use and/or reimbursement and price at national level (countries want to retain control of price as part of their role in controlling healthcare system costs)	92%	4.08	61%	2.92
	2	Convergence of methods to assess costs (facilitated by EunetHTA). Separate decisions on use and/or reimbursement and price at national level	45%	2.77	69%	3.00
	3	There is a move to greater convergence of healthcare systems within the EU. This includes a move to a pan-European assessment (facilitated by EunetHTA) and appraisal of clinical and cost–effectiveness evidence and a single pan-European decision on price and/or reimbursement	7%	1.54	60%	2.92
Demand for postlaunch RE studies by HTA bodies and payers	1	Demand for postlaunch RE studies is low. There is no acceptance that RE data are transferable across countries and little coordination of studies between countries and jurisdictions	53%	2.62	45%	2.46
	2	Demand for postlaunch RE studies rises at country level but there is no transferability and no coordination among countries and jurisdictions. This leads to inconsistent postlaunch data collection by manufacturers, and some replication of studies in different countries	69%	3.15	76%	3.15
	3	Demand for postlaunch RE studies rises and there is greater coordination among HTA bodies/payers (particularly in areas such as orphan drugs) and transferability of data across jurisdictions	61%	3.23	83%	3.69

†1 = Very improbable; 2 = Improbable; 3 = Even odds; 4 = Probable; 5 = Highly probable.
‡1 = Slight-to-nil influence; 2 = Limited influence; 3 = Moderate influence; 4 = Considerable influence; 5 = Dominant influence.
AL: Adaptive licensing; EHR: Electronic health record; EUnetHTA: European Network for health technology assessment; HTA: Health technology assessment; PAES: Postauthorization efficacy study; PASS: Postauthorization safety study; PCT: Pragmatic clinical trial; RE: Relative effectiveness.

Table 3. Round 1 modified Delphi results for key factors (cont.).

Factors	Pathway number	Pathway description	Probability of occurrence (n = 13; scores 3–5)†	Occurrence mean score†	Level of Impact (n = 13; scores 3–5)‡	Level of impact mean score‡
Coordination between regulatory and HTA bodies	1	More coordination and joint requirements for prelaunch and postlaunch studies leading to: coordinated scientific advice for prelaunch studies; coordinated scientific advice for postlaunch studies	60%	2.97	84%	3.54
	2	Coordinated scientific advice for prelaunch studies often leads to agreement on common end points and study design. However, there is no coordination between HTA and regulatory bodies in respect to postlaunch studies (e.g., any 'adaptive reimbursement' requirements are quite separate from any postlaunch AL regulatory studies)	78%	3.08	76%	3.15
	3	Scientific advice for prelaunch studies rarely leads to agreement on common end points and study design. In addition, there is no coordination between HTA and regulatory bodies in respect of postlaunch studies (e.g., any 'adaptive reimbursement' requirements are quite separate from any postlaunch AL regulatory studies)	53%	2.62	46%	2.08
Methodologies to analyze RE evidence	1	Little progress is made on methodological issues relating to methods	29%	2.31	30%	1.92
	2	There is increased debate and more funding supporting methodological research on RE methods. Real progress is made, but no common agreement on which methods should be used and how robust they are	53%	3.54	67%	2.85
	3	Progress and agreement on standard methodologies for many of the elements of RE analysis to support decision-making are achieved	61%	2.85	84%	3.46
	4	Progress and agreement on techniques to facilitate analysis and utilization of big data (large repositories of unstructured or semistructured health information, including laboratory data and information on bio specimens, genomic or biomarker data, among others) are achieved	61%	3.00	84%	3.54

†1 = Very improbable; 2 = Improbable; 3 = Even odds; 4 = Probable; 5 = Highly probable.
‡1 = Slight-to-nil influence; 2 = Limited influence; 3 = Moderate influence; 4 = Considerable influence; 5 = Dominant influence.
AL: Adaptive licensing; EHR: Electronic health record; EUnetHTA: European Network for health technology assessment; HTA: Health technology assessment; PAES: Postauthorization efficacy study; PASS: Postauthorization safety study; PCT: Pragmatic clinical trial; RE: Relative effectiveness.

Table 3. Round 1 modified Delphi results for key factors (cont.).

Factors	Pathway number	Pathway description	Probability of occurrence (n = 13; scores 3–5)†	Occurrence mean score†	Level of Impact (n = 13; scores 3–5)‡	Level of impact mean score‡
Infrastructures to conduct RE research	1	Standardization and implementation of EHRs in many health systems enables access to relevant patient level data (e.g., resource use, interventions, disease progression, outcomes)	61%	2.85	75%	3.69
	2	Disease registries are increasingly used to collect patient level data in many EU countries	99%	3.77	91%	3.92
	3	Observational databases (e.g., the Clinical Practice Research Database in the UK) are increasingly developed to be used for prospective and retrospective observational studies and to support PCTs by enabling randomized patients to be tracked	85%	3.38	92%	3.77
	4	Research infrastructure in primary and secondary/tertiary care are developed to support PCTs	52%	2.77	76%	3.46

†1 = Very improbable; 2 = Improbable; 3 = Even odds; 4 = Probable; 5 = Highly probable.
‡1 = Slight-to-nil influence; 2 = Limited influence; 3 = Moderate influence; 4 = Considerable influence; 5 = Dominant influence.
AL: Adaptive licensing; EHR: Electronic health record; EUnetHTA: European Network for health technology assessment; HTA: Health technology assessment; PAES: Postauthorization efficacy study; PASS: Postauthorization safety study; PCT: Pragmatic clinical trial; RE: Relative effectiveness.

other EU countries. Little progress is made from current levels of pan-EU coordination between HTA bodies or dialogue between the EMA and HTA bodies. Postlaunch studies are requested in some countries, and some progress is made in national use of EHRs and disease registries. However, the emphasis for the EMA and national HTA bodies remains on 'at launch' appraisal with little enthusiasm for entering into postlaunch commitments or imposing requirements which they suspect may not be met by companies. Companies are expected to finance and organize any RE research that is required. As a consequence, we speculate that companies are likely, under this scenario, to concentrate on improving the data they are able to generate prelaunch and to limit postlaunch commitments.

Scenario 2 – some changes; regulatory dominates

Scenario 2 focuses on regulatory change. This highlights the use of PAESs, envisaging greater use of postlaunch studies by the EMA. This complicates the ability of companies to meet potentially conflicting demands for postlaunch data by HTA bodies/payers and regulatory bodies. In reality, this will mean that EMA requests for PAESs are likely to reduce the ability of HTA bodies to require additional studies over and above the PAESs. Conversely, companies are likely to want to agree PAES designs and end points with the EMA that are likely to meet any HTA body requirements for postlaunch studies. Companies will, as in scenario 1, put most emphasis on improving prelaunch data, but in this scenario this is made easier as there is some alignment between the EMA and HTA bodies on their respective requirements, and EUnetHTA has succeeded in both reducing the methods differences between HTA bodies/payers and in reducing duplication of effort, by getting an element of mutual recognition of RE evidence, making it easier for companies to plan meeting prelaunch HTA body requirements. Companies are more willing to invest in postlaunch evidence collection as part of a PAESs but reluctant to add further commitments requiring additional studies.

Scenario 3 – major changes; high trust environment

Scenario 3 is a favorable environment for generating RE data both pre- and post-launch, with greater coordination between regulatory and HTA body/payers, and potential collaboration by health systems in the collection and analysis of data for postlaunch studies. It includes the use of adaptive licensing in a broad set of disease areas. This is assumed to be a high trust environment. Companies can work with regulators and HTA bodies/payers to get the right balance of prelaunch and postlaunch collection of RE data.

The three scenarios, together with the nine key factors driving them, are summarized in Table 4 and set out in more detail in Supplementary Box 1.

The likelihood of the scenarios occurring

The likelihood of the three scenarios occurring was voted on in round 3 of the modified-Dephi. Scenario 2 (some change: regulatory dominates) was regarded as the most likely to occur, with scenario 3 (major change: high trust) second. A continuation of the status quo scenario 1 (no change: low trust) was thought least likely of the three scenarios to occur. The results are set out in Table 5 below.

Thus the majority participating in the final round of the modified Delphi exercise concluded that change was more likely to occur than not. More than half of stakeholders thought that there would be a degree of convergence between HTA bodies and greater dialogue between the EMA and HTA bodies. The EMA would become more involved in commissioning studies postlaunch, although this would stimulate further dialogue between companies, HTA bodies and the EMA around postlaunch RE evidence. Together with some progress in data generation there could be a very different expectation of and ability to supply RE evidence. Nearly half of respondents thought the changes could be major, and more conducive to greater use of RE evidence, with the potential for agreement prelaunch as to the type of pre-and postlaunch evidence to be provided.

Limitations of the study

The scenarios are plausible frameworks for the future generated through a rigorous methodology. Scenario exercises of this type have four main limitations.

First, by definition, the scenarios cannot anticipate sudden unpredictable change, in other words, a discontinuity with existing trends. Were any ruptures to occur, however, the impact they would impact in a 7-year time frame may be limited.

Second, such an exercise does not deal well with 'Wild Cards', in other words, factors and pathways that are identified as having low probability of occurrence, but high impact were they to occur. We found very few factor/pathway combinations that were deemed unlikely to occur and yet expected to be impactful were they to occur. Only one had an impact mean vote above 3.5 and this was incorporated in a scenario at Delphi round 2.

Third, a potential limitation of the modified Delphi approach is that a different group of experts might come to different conclusions. We sought to reduce this type of bias by targeting experts across the key stakeholders and with different skill sets. As a sepa-

Table 4. Summary of final scenarios.

Factors	Scenario 1 – no change; low-trust environment	Scenario 2 – some changes; regulatory dominates	Scenario 3 – major changes; high-trust environment
Regulatory			
Role of PAESs	New pharmacovigilance legislation leads to little change	PAESs implemented in a systematic way	Integrated system with use of PAESs and AL
Use of AL	CMA used sparingly	AL is progressed through more use of CMA	AL applied to a variety of drugs beyond CMA
HTA bodies/payers			
Methods for assessment and sharing of assessments	No agreement on methods for clinical assessment	Methods convergence and clinical assessment sharing across many but not all HTAs	Methods convergence and clinical assessment sharing across many but not all HTAs
Location of decision-making on pricing and use	Decision-making remains national or subnational	Decision-making remains national or subnational	Decision-making remains national or subnational
Demand for postlaunch studies	Postlaunch studies requested in some countries and to be conducted locally	Reduced ability to request local postlaunch studies given role of regulatory PAESs	Coordination of demand for postlaunch studies (often linked to conditional reimbursement schemes)
Dialogue regulatory/HTA	Limited interaction and little alignment pre- or postlaunch	Some alignment in prelaunch requirements	Joint thinking for both pre-and postlaunch requirements
RE research infrastructures, methods, funding			
Improvement in infrastructure for RE studies	Increased use of disease registries and some countries progress EHRs	Increased use of disease registries and some countries progress EHRs	Collaborations across large registries. Many countries progress EHRs. Progress in use of big data
Work on methods	Limited methods development	Limited methods development	Progress in methods development
Funding/commissioning	Industry is responsible for doing RE research	Industry is responsible for doing RE research	Public–private partnerships have a major role

AL: Adaptive licensing; CMA: Conditional marketing authorization; EHR: Electronic health record; HTA: Health technology assessment; PAES: Postauthorization efficacy studies; RE: Relative effectiveness.

Table 5. Results of voting in the modified-Delphi round 3: likelihood of scenarios.

Scenario name	Percentage of votes assigned to each ranking				
	1 – Highly unlikely	2 – Unlikely	3 – Even odds	4 – Likely	5 – Highly likely
Scenario 1: *status quo*; low trust	14	21	36	14	14
Scenario 2: moderate change; regulatory dominates	0	21	21	50	7
Scenario 3: major change; high trust	0	29	29	43	0

rate exercise, postcompletion of the modified Delphi process, we recruited a completely separate group of expert stakeholders not involved in the original future scenarios exercise and asked them to vote on the likelihood of occurrence by 2020 of each of the three scenarios generated by the modified Delphi process. The external review respondents were more inclined to be skeptical about a change in the environment for RE in Europe. However, when we compared voting from the external review, and the voting from the modified-Delphi round 1 on the four initial scenarios, the results were very similar. The effect of the Delphi process was to bring the participants to the conclusion that more change was likely to occur than they had originally expected. The external stakeholders, not exposed to this process, had expectations of likelihood that were similar to those of the Delphi panel at the start of the modified-Delphi process.

Fourth, the modified-Delphi process could introduce a bias whereby a few strong voices lead the participants to develop unrealistic expectations of change in a particular direction. Although the process led participants to have more expectation of change than at the start, the discussions at the Workshop were robust and informative, and there was no evidence of one particular viewpoint dominating others. Our conclusion is that the modified-Delphi worked as it should in bringing a group of experts to a different point of view then that with which they started through a process of interaction and exposure to the work of the project team and the views of others.

Discussion of the results

The particular focus is on 'relative' or 'comparative' benefit in routine clinical use of the product. Although RE evidence can be generated prelaunch by making trials more pragmatic, it is more often thought of as a postlaunch activity, with PCTs and observational studies offering ways of collecting RE evidence.

The key points of the most likely scenario (scenario 2: some change; regulatory dominates) are threefold.

Scenario 2: some change; regulatory dominates

First, greater HTA and EMA coordination prelaunch enables companies to undertake additional investment in prelaunch RE evidence generation, in the expectation of a postive return on investment, in other words, they are more likely to be able to meet EMA and HTA body 'at-launch' evidence requirements. Indicators of progress in this area include the extent to which the EUnetHTA and EMA parallel scientific advice initiatives succeed in achieving a degree of consensus between HTA bodies, the regulator and the company, and evidence that companies following scientific advice achieve quicker access to patients from the regulator and from European payers. Further evidence will come from the success or otherwise of the EUnetHTA pilot REA project currently underway, for the mutual recognition of assessments of prelaunch RE evidence undertaken by designated lead HTA bodies.

Second, the implementation of PAESs turns the EMA into a significant actor in the postlaunch RE environment. This is potentially disruptive. It may reduce company willingness to undertake additional postlaunch RE evidence collection for HTA bodies, particulary in an environment in which there is little cross-border collaboration between HTA bodies as to postlaunch requirements or acceptance of any transferability of the results of PCTs and observational studies from one country to another. The way in which PAESs is implemented will be a key factor impacting on the future environment for RE studies. Indicators of the direction of travel will be the flexibility in study design set out in the EMA scientific guideline; the extent to which it is used in (say) the first year following its introduction; and EMA willingness to involve HTA bodies in decisions as to what types of studies to require.

Third, progress, albeit limited, in the use of EHRs and in the populating of disease registries in some countries offers the prospect of increasing the potential for (and reducing the cost of) conducting postlaunch PCTs and observational studies. Concerns remain around the interpretation of the results of observational

studies. Indicators of progress might be the numbers of studies conducted by research groups for companies and other stakeholders, and changes in the degree of infrastructure challenges such researchers faced, particularly in conducting PCTs. However, in scenario 2 the benefits of even limited additional data sources and infrastructure improvements may not be realized. This is because companies willingness to fund postlaunch studies for HTA bodies is limited, given the potential need to meet regulatory PAES requirements and a low expected payback from undertaking them.

Scenario 3: major changes; high trust environment

In the alternative likely scenario, scenario 3: major changes; high trust environment, the three key points are set out below.

First, coordination takes place across HTA bodies in their demand for postlaunch studies. This could be seen as an evolution of the EUnetHTA and EMA led initiatives to reduce prelaunch duplication of effort through both coordinated scientific dialogue with companies, with mutual recognition by HTA bodies of assessments of clinical evidence. There is also a greater willingness of payers to offer CED arrangements, including performance-based risk sharing arrangements (PBRSAs) [18] to companies enabling products to be reimbursed and available to patients while further RE evidence is collected to address uncertainty. There is recognition that much (but not all) RE evidence is transferable across national boundaries. An indicator of travel would be the numbers of CED/PBRSA arrangements that are put in place, the variety of payers involved and whether any schemes use data collected outside of the local health system? A combination of these trends would involve, for example, one study meeting the requirements of an EMA PAESs and one or more HTA body requirements linked to CED/PBRSA arrangements. This would require EMA dialogue with HTA bodies before PAES designs were finalized. An early indicator of progress in this direction will be whether there is any HTA input into the EMA's deliberations about the nature of any PAES requirement.

Second, the ability to undertake postlaunch RE evidence collection is substantially improved by much greater use of EHRs and pan-EU collaborations on disease registries. More data are available for observational studies. PCTs become easier to conduct as patients can be followed through their EHRs. Improved methods for analyzing confounding issues increase the confidence of HTA bodies and the EMA in observational

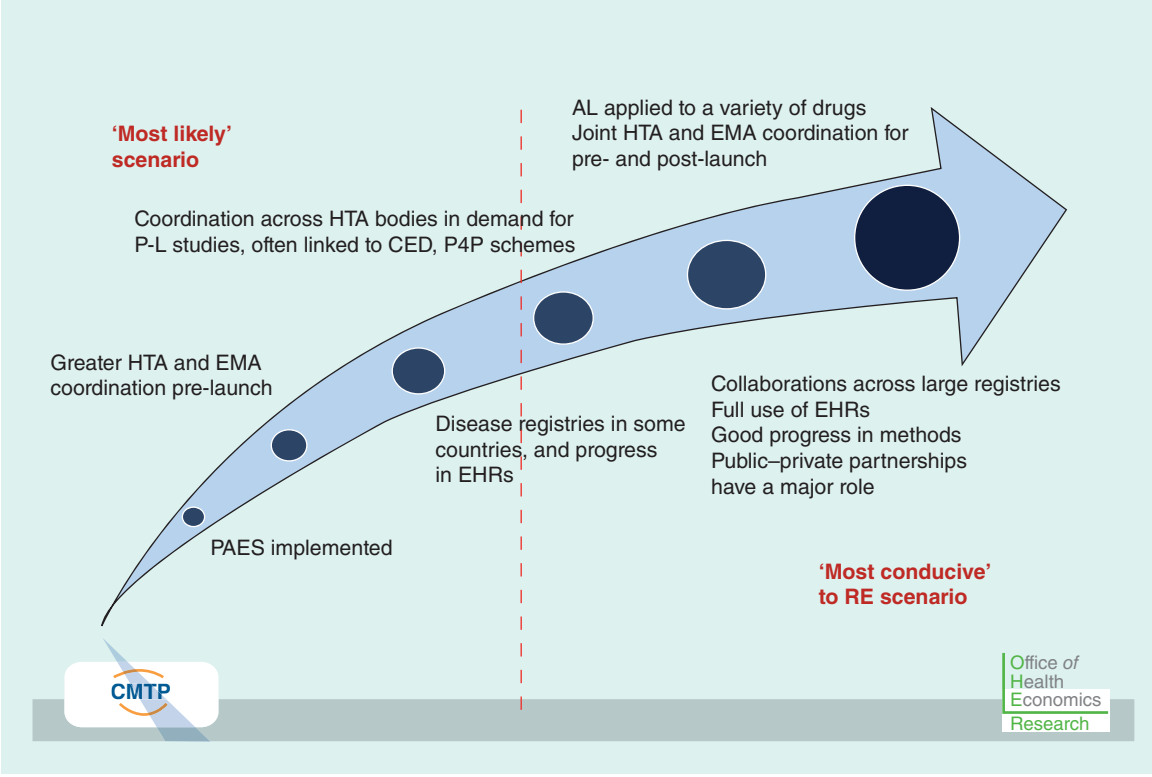

Figure 2. Possible evolution of relative effectiveness evidence requirements in Europe.
CED: Coverage with evidence development; EHR: Electronic health record; HTA: Health technology assessment; PAES: Postauthorization efficacy studies; P-L: Post-launch; RE: Relative effectiveness.

studies, albeit conditioned on a continued initial preference from payers and regulators for randomized evidence. Health systems (payers and providers) and HTA bodies take a more active interest in the generation of RE evidence, and public–private partnerships develop to organize and finance studies. Indicators of progress in this direction will be the numbers of observational studies that use crossborder disease registries; the numbers of PCTs conducted; and the numbers of studies conducted by public–private partnerships.

Third, AL is widely applied in the EU (beyond the remit of CMA) with PAESs providing a tool for the EMA to address uncertainties at initial market authorization. Dialogue between the EMA, HTA bodies and companies develops in a way that enables prelaunch scientific advice to address trade-offs between greater RE evidence collection pre- versus post-launch. Companies can then choose the drug development route that meets EMA and HTA body requirements in the most efficient way while making the drug available for patients and earning a return on their investment. An early indication of progress in this area will be the success or otherwise of the EMA call to industry to put forward 'assets' for AL pilot [7]. A further indicator will be whether there is any HTA input into the EMA's deliberations about the nature of any PAES requirement.

It is helpful to think of the potential evolution of the demand for, and ability to supply, RE evidence as a journey over time. Which point in that journey may have been arrived at in 2020 remains to be seen, but the difference between the two scenarios of change beyond the status quo can be characterized by reference to key points of evolution. We have summarized a possible evolution in Figure 2.

Stakeholders wishing to understand which, if any, of the three scenarios appears to be unfolding over time can conduct an audit of the environment using these key points of reference highlighted above.

Conclusion

The purpose of the scenario exercise was to better understand the demand for, and ability to generate, RE evidence for drugs in Europe in 2020.

Participants concluded change in the environment set out in scenario 2 was most likely to occur. The EMA becomes more involved in commissioning postlaunch studies, stimulating some further dialogue, but not agreement, between companies, HTA bodies and the EMA around postlaunch RE evidence. There is some convergence on prelaunch requirements between the EMA and HTA bodies. Together with progress in EHRs and pan-European disease registries, there could be very different expectations of, and ability to supply, RE evidence. Many respondents thought the changes could be even more conducive to greater use of RE evidence (as set out in scenario 3) with major changes, including use of adaptive licensing by the EMA, and the use of conditional approvals by HTA bodies and payers and greater convergence of EMA and HTA body on both pre- and post-launch evidence requirements. We note also, however, that fiscal and other constraints may produce a 'status quo' scenario in which there is little trust between the parties, and little change from today. How quickly Europe moves beyond the status quo, along the spectrum toward the scenarios of change, remains to be seen.

Future perspective

According to the exercise described in this report, the next 7 years or so in Europe are likely to be characterized by moves toward regulatory reform, greater pan-European coordination of HTA performed by HTA bodies on behalf of payers, and greater dialogue between regulatory and HTA bodies. Specifically, the EMA will become more involved in commissioning efficacy/effectiveness studies postmarketing authorization (postlaunch), which will stimulate further dialogue between companies, HTA bodies and the EMA around postlaunch RE evidence requirements. There will be a degree of convergence between HTA bodies on at-launch evidence requirements and assessment processes and greater dialogue between the EMA and HTA bodies around prelaunch parallel scientific advice for companies. Together with some progress in availability of EHRs in national health systems and in the development of pan-European disease registries there will be a very different expectation of, and ability to supply, RE evidence both prelaunch and postlaunch. Drug makers will respond by building elements of RE into the traditional drug development paradigm, exploring the potential for incorporating more pragmatic approaches to clinical trial design pre- and post-launch, and for using postlaunch observational studies to meet outstanding regulator and HTA body requirements for RE evidence. The initial emphasis will be on improving prelaunch evidence, but will be followed by a greater emphasis on improving the quality of postlaunch evidence.

Acknowledgements

The authors are grateful for the 27 experts who gave their time to participate at various stages of the modified Delphi exercise.

Financial & competing interests disclosure

This study was funded by research grants from Amgen, GSK, Lilly, Novartis and Sanofi. The authors are grateful for their funding. A Towse and M Garau work for the Office of Health Economics which receives research and consulting income

from the ABPI and from individual pharmaceutical companies. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct

The authors state that they have obtained appropriate institutional review board approval or have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations. In addition, for investigations involving human subjects, informed consent has been obtained from the participants involved.

Executive summary

- Patients and those paying for healthcare in the EU (Europe) increasingly want evidence of the additional benefit one treatment offers as compared with another in routine clinical use – termed relative effectiveness (RE) in Europe.
- This exercise uses a robust methodology, involving all the main stakeholders, to understand the interaction between the factors influencing the future environment for generating and using this evidence in Europe.
- Using semistructured key informant interviews and three rounds of a modified Delphi approach, we gathered expert perspectives on key factors influencing the future development of RE evidence.
- The most influential factors were the degree of regulator use of postauthorization efficacy studies and of adaptive licensing; health technology assessment (HTA) body coordination across Europe; regulator–HTA interaction both before and after drug marketing authorization (pre- and post-launch).
- We combined different pathways for these factors into three plausible scenarios exploring possible future demand for, and ability to produce, evidence of the RE of a new drug by the year 2020. The scenario voted most likely, 'some change: regulatory dominates,' entailed the EMA becoming more involved in commissioning studies postlaunch.
- This would stimulate further dialogue but not necessarily agreement between companies, HTA bodies and the EMA around postlaunch RE evidence. There would be a degree of convergence between HTA bodies and greater dialogue between the EMA and HTA bodies on prelaunch requirements.
- Together with some progress in availability of electronic health records and pan-European disease registries there could be a very different expectation of and ability to supply RE evidence.
- Nearly half of respondents thought the changes could be even more conducive to greater use of RE evidence (as set out in scenario 3: major change: high trust), with the potential for agreement prelaunch as to the type of pre- and post-launch evidence to be provided, combining the use of adaptive licensing by the EMA with the use of coverage with evidence development agreements by HTA bodies and payers.
- It is helpful to think of the potential evolution of the demand for, and ability to supply, RE evidence as a journey. Where we have got to by 2020, along the spectrum between the two scenarios of change, remains to be seen.

References

1 High Level Pharmaceutical Forum. Final Conclusions and Recommendations of the High Level Pharmaceutical Forum. http://ec.europa.eu/enterprise/sectors/healthcare/files/docs

2 EUnetHTA. Joint Action 2, 2012–2015. www.eunethta.eu/activities/EUnetHTA%20Joint

3 European Medicines Agency. Road Map to 2015. www.ema.europa.eu/docs/en_GB/document_library

4 European Union. Commission Delegated Regulation (EU) No. 357/2014. Situations in which post-authorisation efficacy studies may be required. http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri

5 European Commission. Post-authorisation efficacy studies. Consultation document. http://ec.europa.eu/health/files/pharmacovigilance/2012

6 Boon WPC, Bouvy JC, Broekmans AW *et al.* Improving the EU system for the marketing authorisation of medicines. Learning from regulatory practice. http://escher.tipharma.com/fileadmin/media-archive/escher

7 European Medicines Agency. Adaptive licensing pilot project. www.ema.europa.eu/ema/index.jsp?curl=pages/regulation

8 Eichler H-G, Oye K, Baird LG *et al.* Adaptive licensing: taking the next step in the evolution of drug approval. *Clin. Pharmacol. Ther.* 91(3), 426–437 (2012).

9 Baird LG, Banken R, Eichler H-G *et al.* Accelerated access to innovative medicines for patients in need. *Clin. Pharmacol. Ther.* 96(5), 559–571 (2014).

10 Eichler H-G, Baird LG, Barker R *et al.* From adaptive licensing to adaptive pathways: delivering a flexible life-span approach to bring new drugs to patients. *Clin. Pharmacol. Ther.* 97(3), 234–246 (2015).

11 Innovative Medicines Initiative. GETREAL: incorporating real-life clinical data into drug development. www.imi.europa.eu/content/getreal

12 Trusheim M. Better science, better health. New pathways and new sources of evidence, what do we need? http://vitaltransformation.com/wp-content/uploads

13 Towse A. New challenges for improving european health care: earlier access to innovation – is it worth it? Presented at:

ISPOR 17th Annual European Congress, Amsterdam, The Netherlands, 8–12 November 2014.

14 Baird LG, Trusheim MR, Eichler H-G *et al.* Comparison of stakeholder metrics for traditional and adaptive development and licensing approaches to drug development. *Ther. Innov. Regulat. Sci.* 47(4), 474–483 (2013).

15 Messner D, Mohr P, Towse A. Futurescapes: evidence expectations in the US for CER for drugs in 2020. *J. Comp. Eff. Res.* doi:10.2217/CER.15.6 (2015) (Epub ahead of print).

16 Messner D, Towse A, Mohr P, Garau M. The future of evidence generation for comparative effectiveness and relative efficacy of drugs, an international perspective. *J. Comp. Eff. Res.* doi:10.2217/CER.15.8 (2015) (Epub ahead of print).

17 EunetHTA. The final version of HTA Core Model® and the Methodological Guidelines for Rapid REA of Pharmaceuticals. www.eunethta.eu/news/final-version-hta-core-model-and

18 Garrison, LP Jr, Towse A, Briggs A *et al.* Performance-based risk-sharing arrangements – good practices for design, implementation, and evaluation: ISPOR good practices for performance-based risk-sharing arrangements task force report. *Value Health* 16, 703–719 (2013).