

GetReal - Project No. 115546

**WP1: Deliverable D1.6 (Case Study Review: Rheumatoid
Arthritis)**

Lead Organisations and Investigators:

University of Leicester,

Mats Ericson (Amgen), Sandro Gsteiger (Roche)

Executive Summary

This report describes the outcomes of stakeholder engagement undertaken by the GetReal Work Package 1 of which had the aim of eliciting a comprehensive stakeholder view on the usefulness and acceptability of different analytical approaches to synthesise clinical evidence—including real-world evidence (RWE)—to establish the relative effectiveness of new drugs in rheumatoid arthritis (RA). Stakeholder opinions were elicited during a workshop that took place on Wednesday March 16th, 2016 in London, UK.

Workshop objectives

The primary aim of the workshop was to facilitate stakeholder discussions on the potential role of RWE, in particular registry data, in overcoming key effectiveness challenges for the assessment of biologics in RA, by exploring new evidence synthesis methods and their potential impact on decision making, as well as, their generalisability to other disease areas.

Stakeholder feedback was solicited from the perspective of Pharma R&D, regulators, HTA bodies, patients and practitioners. Methods and results for each of the analytical approaches listed below and detailed in the WP1: Deliverable 1.5 (Case Study Review: Rheumatoid Arthritis) were presented to stakeholders.

Six analytical approaches were considered in the RA case study to address two key issues: connecting disconnected second-line networks of evidence with RWE and bridging evidence between first- and second-line networks with RWE:

Approach 1: 'naïve' pooling

Approach 2: exchangeability assumption

Approach 3: RWE approach (SCQM¹ registry data)

Approach 4: RWE approach (SCQM and BSRBR² registry data)

Approach 5: univariate network meta-analysis (NMA)

Approach 6 (a/b/c): bivariate NMA (under different assumptions/correlation structures)

¹ Swiss Clinical Quality Management in Rheumatic Diseases registry

² British Society for Rheumatology Biologics Registers (Rheumatoid Arthritis Register)

Workshop participants were asked to consider the following questions during break-out sessions and group discussions focusing on the analytical methods highlighted in the workshop:

- Under what circumstances would these methods have an impact on decision-making?
- Could these methods be used to address other 'effectiveness' issues and/or applied to other disease areas?
- How could these methods be improved and developed further to be useful and acceptable to stakeholders?
- What issues might stand in the way of adopting such methods by various stakeholders?
- Are there situations where these methods would be particularly useful (or not useful)?
- How can we best communicate the implications of these methods to engage a broader range of stakeholders?

Conclusions

Overall, stakeholders were interested and inquisitive about how the analytical approaches presented could reduce decision-making uncertainty by including RWE. However, it was felt that HTA bodies were most likely to benefit from these approaches since their overarching aim was to produce relative effectiveness estimates by extending existing evidence synthesis methods (i.e. NMA). Nonetheless, specifically regarding the approach 5 and 6, the techniques presented to bridge networks of evidence could 'in principle' also be used by Pharma R&D to better understand the gaps in the evidence base across lines of therapy and to aid in the design of future clinical trials. The robustness of these methods was a key acceptability factor for HTA bodies/Pharma R&D and all stakeholders agreed that analyses adding further decision uncertainty would not be deemed helpful at any stage of the product development cycle.

Several concerns were raised around potential biases that could be introduced by making strong assumptions and/or including RWE in the different approaches put forward. Concerns expressed by workshop participants included issues associated with data quality, methodology and practical issues. For example, on data quality, the two main issues raised were the selection of the RCTs and RWE studies and the introduction of bias and confounding when using registry data in NMA (e.g. Approaches 3 and 4). It was noted that increased heterogeneity in the evidence base and potential inconsistencies between randomised and registry data may widen the credible intervals for the relative effect estimates, thus leading to a 'counter-productive' approach and rendering the proposed methods of limited use or impact.

The strength of the assumptions made and their plausibility was extensively discussed during the workshop. It was felt that more justification of these assumptions was required and that clinical input from both practitioners and registry holders would be beneficial. Participants were in agreement that further demonstration of the robustness of the approaches was required as large uncertainty in the results presented, as well as the sometimes large variation in the point estimates, raised concerns about the methodology suggesting underlying data limitations (i.e. inconsistencies within the evidence base and across networks).

The complexity of the methods was also an issue; it was felt that 'breaking down' these approaches in more manageable and acceptable analyses could be more constructive in the short-run (for instance, analyses generated for validation purposes, for estimating correlation structures but not necessarily extending that to predicting relative effectiveness).

Lastly, participants concluded that more efforts should be directed towards improving the credibility of RWE overall to ultimately improve the confidence amongst all stakeholders.

Table of contents

Executive Summary	2
Workshop objectives	2
Conclusions	3
Table of contents	5
1 Decision-making perspectives	6
2 Stakeholder views – Approaches 1-4	7
2.1 Pharma R&D	7
2.2 Regulators	8
2.3 HTA bodies	8
2.4 Patients and Practitioners	10
3 Stakeholder views – Approaches 5-6	12
3.1 Pharma R&D	12
3.2 Regulators	13
3.3 HTA bodies	13
3.4 Patients and Practitioners	14
4 Concluding remarks	17
4.1 RA specific observations on the use of RWE:	17
4.2 Recommendations related to methods (network meta-analysis):	17
4.3 Recommendations regarding use of RWE:	18

1 Decision-making perspectives

Up to four decision-making perspectives were considered for each analytical approach:

Pharma R&D who make decisions on medicine development including design and funding of evidence generation programmes and constituent studies.	Regulatory organisations at country and above country-level, who provide marketing authorisation based on evidence of product quality, safety and efficacy.	HTA organisations who advise, and in some cases make decisions on reimbursement of new medicines, based on evidence of safety, efficacy and in particular the (relative) effectiveness of new medicines for their populations.	Patients and practitioners who advise, in their own capacity, on the 'real-life' clinical and practical concerns of both drug development and drug use.
-------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------

In order to encourage openness and sharing of information, the workshop was held under Chatham House rules. However, without revealing the identity nor the affiliation of participants, it was agreed the below list of stakeholder groups represented at the workshop could be shared:

Academia	University of Bern, University of Leicester, University of Manchester, University of York
Pharma R&D	Amgen, Janssen, Lilly, Roche, Sanofi
Regulators and HTA Bodies	EMA (European Medicines Agency), NICE (National Institute for Health and Care Excellence), ZINL (Zorginstituut Nederland)
Patients and practitioners	IAPO (International Alliance of Patients' Organizations)
Registries/RWE representatives	BSRBR (British Society For Rheumatology Biologics Register), RABBIT (Rheumatoide Arthritis: Beobachtung der Biologika-Therapie (German register))
Other institutions	Evidera, OHE (Office of Health Economics)

2 Stakeholder views – Approaches 1-4

Analytical approaches 1 - 4: Connecting disconnected networks of evidence with RWE to generate estimates of relative efficacy/effectiveness

In disease areas where numerous treatments are available for a single indication, decision makers must rely on numerous sources of information when undertaking approval/appraisal decisions. In this regard robust synthesis of available data sources is important. These approaches explored the role for RWE in connecting otherwise disconnected networks of evidence in order to conduct NMA and obtain relative efficacy/effectiveness estimates for all relevant treatment comparisons.

Comparing 'naïve' methods of pooling to NMAs using RWE to connect networks of evidence allowed us to calculate relative effects otherwise not attainable and demonstrated an increased precision in the treatment estimates compared to the base case, as well as, more conservative point estimates.

2.1 Pharma R&D

Usefulness

It was mentioned that while the approach may have some intrinsic utility for Pharma R&D, for example when a treatment comparison is required by HTA bodies but the network of evidence is disconnected; however, this is beyond the scope of the decision making that regulators would undertake during the review of a regulatory submission.

Situations

Participants suggested the specific situation where such methods may be considered by Pharma R&D is if HTA bodies required the assessment of a treatment comparison not evaluated in a head-to-head RCT and not part of a closed or connected network of evidence. The use of these methods is contingent on their acceptability by decision-makers; therefore, it was felt by participants that more work would be required to justify the assumptions underpinning the approaches, to ensure the RWE is adequately identified and accounted for in the analysis, and to appropriately assess heterogeneity, inconsistency and uncertainty. In addition, it was thought that these approaches should air on the side of caution and more conservative assumptions would be more likely to be accepted by HTA bodies.

Issues

'Naïve' pooling was regarded as the most simplistic approach whilst the exchangeability assumption effectively places a class effect on treatments and was thought to be more justifiable. However, both approaches require strong assumptions and exchangeability is untestable, a more in-depth review of the

evidence—randomised and observational—would be required to compare patient and trial characteristics, treatment effects and drug pathways, and to establish plausibility of assumption.

Registry data was acknowledged as a valuable source of evidence. Matching registry data to the trial inclusion criteria and choice of outcome did restrict the number of observed patients from the SCQM and BSRBR significantly.

Communication

Overall, greater transparency of the methodology is needed in particular when selecting randomised and observational evidence and when making modelling assumptions (such as prior distributions, exchangeability). In addition, more clarification would be needed to justify when it would be relevant and appropriate to use such these approaches.

2.2 Regulators

Usefulness

Participants noted that the decision making process of regulators is different from that of Pharma R&D developers and HTA agencies. For this reason there were some reservations about how useful this approach would be in a regulatory setting.

Situations

Not applicable

Issues

There was concern that in general the approach relied on assumptions that were unverifiable and additionally would be subject to the same biases generally associated with RWE studies. This in turn would increase uncertainty rather than reduce it.

Communication

Not applicable

2.3 HTA bodies

Usefulness

Participants expressed the view that NMAs have a role in disease areas with multiple comparators where head-to-head treatment comparisons are not available. The approach of incorporating RWE in data synthesis was supported by some participants, such as in this instance, where limited RCT data are available and a disconnected network of evidence does not permit multiple treatment comparisons.

In addition, it was felt by the majority of stakeholders present at the workshop that HTA offered the most appropriate context to use these methods and HTA bodies

were most likely to benefit from these approaches. However, not all HTA bodies agreed that RWE should be used (even as supplementary data) and cautioned that the acceptability or generalisability of these methods across HTA bodies may be challenging.

Situations

The approaches could be useful in specific situations where there are no direct or indirect RCT data available to compare treatments of interest for decision-making, i.e. the network of randomised evidence is disconnected and does not permit NMA to estimate all relevant relative treatment effects.

Issues

Issues raised by participants concerned data quality, methodology and practical issues. On data quality, the two main concerns expressed by participants was on the selection of the RCT and RWE and the use of registry data introducing bias and confounding in the NMA. Increased heterogeneity in the evidence base and potential inconsistencies between the randomised and registry data may widen the credible intervals for relative effect estimates, leading to a limited value for decision making or rendering the approach unusable. For example, in the case of RA, inconsistency could be driving the reduction in uncertainty and the change in the point estimates from base case to approaches 3 and 4.

Appropriate endpoints in RWE also need to be available in order for these approaches to be viable. The challenges of using DAS remission based on the reporting of this outcome in RCTs and registry data have been discussed in the workshop summary (cf. “1. RA specific observations on the use of RWE”) but may have been inadequate and led to vague results.

On the issue of methodology, participants felt that transparency was needed regarding the methodology and the assessment of its appropriateness for use in a particular situation. Namely, the strength of the assumptions made and their plausibility was extensively discussed by participants. It was felt that more justification of these assumptions was required and that clinical input from both practitioners and registry holders would be beneficial at this stage.

The approach may also be limited by practical issues, such as the additional time required to conduct the data synthesis and for HTA bodies to scrutinise the application of the method and the results. In any case, the impact of introducing RWE in this way should be assessed by means of sensitivity analyses, which ideally should be pre-specified. Participants were in agreement that further demonstration of the robustness of the approaches is required.

Communication

It was felt that the approach has to be made more ‘user-friendly’ as it is too technical and not sufficiently transparent. More emphasis should also be placed on explicitly

indicating how RWE contributes to the effectiveness estimate as part of the expanded NMA. Diagrams may also facilitate the interpretation of results; however, it was noted that in this instance forest plots and ranking charts were not informative or feasible given the large uncertainty in the DAS remission odds ratios.

2.4 Patients and Practitioners

Usefulness

In terms of impact on the wider set of stakeholders, it was felt that the approach has some potential to provide patients and clinicians with additional information on effectiveness.

Situations

Not applicable

Issues

In addition to the issues shared with other stakeholders, there were some concerns expressed regarding the difficulties 'in practice' of classifying trials and registry data in terms of lines of therapy. Clinical input should be sought when handling such complex evidence base, and expert input on how to handle registry data from registry holders would be helpful to optimise the design of these analyses.

Communication

In order for key messages and conclusions from such approaches to be understood and embraced by patients and practitioners alike, methods and results should be communicated more clearly to a wider non-technical audience. The value of generalising methods to incorporate RWE should also be weighed against more patient-centred or individualised approaches.

Table 1 summarises the stakeholder responses for analytical approaches 1-4: connecting disconnected networks of evidence using RWE for NMA to generate relative effectiveness estimates.

Table 1: Stakeholder responses regarding connecting disconnected networks of evidence using RWE for NMA to generate relative effectiveness estimates

	<i>Pharma R&D</i>	<i>Regulators</i>	<i>HTA bodies</i>
Useful in decision making?	Maybe, dependence on acceptability from the other stakeholders	Unlikely	Cautious 'Yes', as long as the addition of RWE into the NMA doesn't increase decision uncertainty, robustness of the methodology is established; and sensitivity analyses are conducted.
Situations	Specific scenario where a comparison with a treatment in disconnected network of evidence is required by regulators/HTA bodies.		Specific scenario where a comparison with a treatment not connected to the available network of evidence is required for the assessment of a new technology.
Issues/ limitations	<p><i>Methodology:</i> assumptions need to be explicit and plausible, and should be conservative.</p> <p><i>Practical:</i> only worth undertaking such approaches if considered 'acceptable' to decision-makers</p>	Dependence on untestable assumptions; bias in RWE.	<p><i>Methodology:</i> the approach is relatively untested and needs further validation</p> <p><i>Data:</i> quality of RWE remains concern and the introduction of bias could be an issue; transparency is needed both in selection of studies/data/outcome and the application of the analytical methods</p> <p><i>Practical:</i> more sensitivity and scenario analyses required to establish robustness of methods</p>
Communication/ adoption	Improve transparency of methodology; ensure assumptions are explicit and plausible.	Improve transparency of methodology; ensure assumptions are explicit and plausible.	Improve transparency of methodology and 'user-friendliness'; clarify and explain underlying assumptions required by each approach, improve communication of both methods and results to non-technical audience.

3 Stakeholder views – Approaches 5-6

Approaches 5-6: Use of NMA including real-world evidence to inform second-line effectiveness estimates from first-line networks of evidence

In RA, the majority of RCTs evaluate treatments in first-line. These approaches attempted to borrow strengths from all the evidence available in the RA, in both first- and second-line, to estimate relative effectiveness estimates.

Two sets of analyses were performed to model the first- and second-line data (from both RCTs and RWE) jointly. First, a univariate NMA was performed for DAS remission at 6 months using data from both registries—SCQM and BSRBR—to bridge evidence between the first- and second-line networks. Second, a series of bivariate NMAs were undertaken assuming both lines of therapy were correlated. The RWE from the registries was used to obtain a correlation estimate between treatment effects in the first- and second-line therapy, as well as, a between-studies correlation between the lines of therapy.

3.1 Pharma R&D

Usefulness

The approaches presented could ‘in principle’ be used by Pharma R&D to better understand the gaps in the evidence base across lines of therapy and to aid in the design of future clinical trials.

Situations

Several situations were cited during group discussion in which this approach could perform better:

- When more RCT and RWE data is available, not just more studies but more data points for comparison;
- When a higher correlation between networks of evidence is observed;
- For more sensitive outcome measures with higher event rates, potentially in other disease areas.

Issues

It was felt that the large uncertainty in the results presented for Analysis 5 and 6, as well as, the large variation in the point estimates raised concerns about the methodology suggesting inconsistency within the networks of evidence. In RA there are differences in the endpoints and the reporting of remission between RWE and RCT studies. As previously mentioned the choice of outcome and scaling issues

could also have influenced the results and clouded the potential applicability of these approaches.

Similarly, the difficulties of comparing DMARDs in the registry data and the ‘control’ arms in the RCTs were discussed. The exchangeability assumption appears to be driving the results more than the correlation between networks of evidence, thus limiting the borrowing of strength and the improvements in uncertainty. These issues challenge the usefulness of RWE in this context.

Communication

It was acknowledged that such an approach could add value to early development planning and has potential to address issues in trial design other than relative effectiveness. However, clear guidance on how these methods would be appraised by HTA bodies would be helpful.

3.2 Regulators

Usefulness

Based on this case study and group discussions, the usefulness of the approaches to bridge first- and second-line evidence in RA was felt to be limited for regulators. The key to deciding whether an approach is useful to regulators and developers should be driven by whether it helps to provide sufficient efficacy and safety information on which to base marketing authorisation; for RA this remains line-specific.

Situations

Not applicable

Issues

Not applicable

Communication

It was felt that more work would be required before the results of this type of approach could be beneficial in a regulatory decision-making context, and importantly the process would have to be transparent with all the assumptions stated clearly.

3.3 HTA bodies

Usefulness

Participants agreed that clearer research questions were needed to ensure that the methods were indeed developed in order to address an issue faced by stakeholders.

However, it was felt these approaches had potential in terms of addressing evidence gaps and maximising a limited evidence base to assess relative effectiveness.

Situations

Specific scenarios where this approach could have been used:

- To compare different subgroups of interest rather than lines of therapy (e.g. pediatric vs. adult population in RA or other disease areas)
- To assess regional differences and borrow strengths across regions/centres

Issues

The main limitation raised was a general comment on the complexity of the methods presented, particularly with regards to their relevance to the research questions being addressed and the available data.

Communication

It was thought that early dialogue among stakeholders on how the approach could best be implemented in study planning and design would be essential in order for value to be added further downstream in the development programme i.e. in regulatory and HTA submissions.

3.4 Patients and Practitioners

Usefulness

Some work has already been done by the BSRBR to evaluate differences in effects between first- and second-line therapies. However, it was felt that the main research question from the perspective of patients and practitioners is how to choose the best treatment option after a treatment failure (DMARD or biologic)? This issue of treatment sequencing was not addressed by the approaches presented.

Situations

Not applicable

Issues

Some of the assumptions made to bridge both networks of evidence were not clinically justified and data was not made comparable across lines of therapy but taken as is, some form of data adjustment would be warranted when comparing treatment (e.g. propensity score matching, baseline risk adjustment, matching by entry year in addition to other covariates).

Sequence of treatments and different treatment pathways is not explicitly considered but such analysis could influence the approach, the design of the analysis, and the conclusions reached in the particular case study.

Communication

It should be acknowledged that some of the issues raised regarding registry data are intrinsic to their nature, i.e. patient registries were initially set-up to collect safety data, and although they capture a wealth of information caution should be used when using the data for effectiveness analyses. Overall, more explanation is needed to clearly define what was done and for what purpose.

Table 2 summarises the stakeholder responses for analytical approaches 5-6: including RWE to inform second-line effectiveness estimates from first-line networks of evidence.

Table 2: Stakeholder responses regarding including RWE to inform second-line effectiveness estimates from first-line networks of evidence

	<i>Pharma R&D</i>	<i>Regulators</i>	<i>HTA bodies</i>
Useful in decision making?	Maybe	Unlikely	Maybe
Situations	There is a willingness to try new methods pre- and post-launch to predict and estimate relative effectiveness estimates; however, internal and external acceptability is an issue.		There is potential for such methods to be useful in HTA, but difficult to generalise approaches from this case study as analyses rely heavily on assumptions and illustrative example did not unequivocally reduce decision uncertainty.
Issues / limitations	Comparability of first- and second-line evidence. Issues with registry data availability and who would ultimately be responsible for conducting such analysis.		Complexity of methods should be addressed to facilitate use and uptake of these methods, if only for assessors to understand/ critique.
Communication / adoption	Improve transparency of methodology; ensure assumptions are explicit and plausible.		Improve transparency of methodology; ensure assumptions are explicit and plausible.

4 Concluding remarks

4.1 RA specific observations on the use of RWE:

RA is a chronic inflammatory disease for which patients will be treated by multiple subsequent therapies over their lifetime. One key issue regarding the evidence base in RA was how RCTs were classified according to line of therapy, as a judgement call may be required to divide trials with mixed population and/or background therapy. For this reason, clearer description of inclusion and exclusion criteria for RCTs is required, especially for populations of interest, to ensure all relevant trials have been identified and adequately divided in first- and second-line of therapies.

Choice of outcome was also a key discussion point amongst workshop participants. The present analysis used DAS remission (DAS<2.6) as endpoint. Remission rates were relatively low under the control intervention MTX, which complicates the estimation of treatment effects and made the interpretation of some results difficult. DAS remission (DAS<2.6) was thought to be a more suitable outcome for MTX-naïve or first-line patients, as patients are less likely to respond and achieve remission with the comparator intervention MTX in second-line than with a new biologic. Unfortunately, few outcomes were commonly reported by both RCTs and RA registries; for example, ACR20 or ACR50 provided the most complete networks of evidence for the trial data but are not measured in the SCQM and BSRBR registries. On the other hand, actual DAS or DAS28 measurements may have been a more appropriate registry endpoint; DAS low disease activity state (DAS<3.2) is more frequently achieved than DAS remission and is also the usual target in routine daily practice. However, IPD may be needed to calculate states not reported in trial publications. These analyses use RWE already available in the form of extensive patient registries in many countries where RA prevalence is high. For these approaches to be feasible, the differences in endpoint definition and the measurement of RA remission over time, as observed between RCTs and RWE studies, would have to be resolved.

The importance of appropriately accounting for comparator intervention was also raised during the group discussion. It was felt that the use of MTX as a control in second-line therapy may not be clinically relevant.

4.2 Recommendations related to methods (network meta-analysis):

NMA would be especially useful for deriving effectiveness estimates in disease areas in which multiple comparators are available but head-to-head evidence is lacking. The acceptability of such methodology would depend on whether issues with NMA in general (study selection, heterogeneity, inconsistency, etc.) would be compounded

by including RWE. It is also conceivable that these RWE-incorporating strategies could be adopted for technologies in different disease areas.

The use of RWE in this manner was thought to be more useful for the relative effectiveness assessment at launch when there may be insufficient direct and indirect evidence on specific treatment comparison of interest. From a regulatory perspective the methodology would likely contribute little information to regulatory submissions. At the workshop, it was felt that the stakeholders most likely to benefit from these approaches were pharma R&D and HTA bodies. However, there was no consensus from HTA organisations and manufacturers in attendance with regards to how best to or under which specific conditions the potential incorporation of the methodology presented would be warranted in their decision making processes.

More emphasis should also be placed on explicitly indicating how RWE contributes to the effectiveness estimate as part of the extended NMA. Unfortunately, the complexity of some of the methods presented in the workshop may impede the uptake and dissemination of these approaches in decision-making. It remains currently unclear to what extent the additional modelling assumptions needed influenced effectiveness estimates. Thorough simulation studies might shed light on the relative impact of modelling assumptions compared to the additional RWE on the results.

4.3 Recommendations regarding use of RWE:

Importantly, RWE would only facilitate stakeholder decision-making and reduce uncertainty if the robustness of the methodology used was established and sensitivity analyses were conducted. All stakeholders indicated that the quality of the RWE ultimately used in analysis remains a concern and the introduction of biases could be an issue; therefore, transparency is needed both in selection of studies/data and the application of any analytical methods.

It was felt throughout the workshop that clearer communication should be prioritised when disseminating the methods and results from this case study to reach both a technical and non-technical audience. In particular, the rationale for undertaking such analyses needs to be stated more strongly, the assumptions made should be further justified, the methods presented should not only be statistically rigorous but clinically relevant, and lastly the uncertainty in the findings should be interpreted more carefully.

Ethics: Do you consider the deliverable is in compliance with the GetReal Ethics section in DoW

Yes

- No (if not please add comments):
- Not applicable
- No: major changes needed, please comment (re-review required)