Quantitative evidence synthesis methods for the assessment of the effectiveness of treatment sequences for clinical and economic decision-making: a review and taxonomy of simplifying assumptions

Lewis, Ruth; Hughes, Dyfrig; Sutton, AJ; Wilkinson, Clare

PharmacoEconomics

Cyswllt i'r cyhoeddiad / Link to publication

25. Apr. 2024

REVIEW ARTICLE

Quantitative Evidence Synthesis Methods for the Assessment of the Effectiveness of Treatment Sequences for Clinical and Economic Decision Making: A Review and Taxonomy of Simplifying Assumptions

Ruth A. Lewis[1] [ORCID] · Dyfrig Hughes[2] · Alex J. Sutton[3] · Clare Wilkinson[4]

Abstract

Sequential use of alternative treatments for chronic conditions represents a complex intervention pathway; previous treatment and patient characteristics affect both the choice and effectiveness of subsequent treatments. This paper critically explores the methods for quantitative evidence synthesis of the effectiveness of sequential treatment options within a health technology assessment (HTA) or similar process. It covers methods for developing summary estimates of clinical effectiveness or the clinical inputs for the cost-effectiveness assessment and can encompass any disease condition. A comprehensive review of current approaches is presented, which considers meta-analytic methods for assessing the clinical effectiveness of treatment sequences and decision-analytic modelling approaches used to evaluate the effectiveness of treatment sequences. Estimating the effectiveness of a sequence of treatments is not straightforward or trivial and is severely hampered by the limitations of the evidence base. Randomised controlled trials (RCTs) of sequences were often absent or very limited. In the absence of sufficient RCTs of whole sequences, there is no single best way to evaluate treatment sequences; however, some approaches could be re-used or adapted, sharing ideas across different disease conditions. Each has advantages and disadvantages, and is influenced by the evidence available, extent of treatment sequences (number of treatment lines or permutations), and complexity of the decision problem. Due to the scarcity of data, modelling studies applied simplifying assumptions to data on discrete treatments. A taxonomy for all possible assumptions was developed, providing a unique resource to aid the critique of existing decision-analytic models.

✉ Ruth A. Lewis
   r.lewis@bangor.ac.uk

1   North Wales Centre for Primary Care Research, College of Health and Behavioural Sciences, Bangor University, CAMBRIAN 2, Wrexham Technology Park, Wrexham LL13 7YP, UK

2   Centre for Health Economics and Medicines Evaluation, Bangor University, Bangor, UK

3   Department of Health Sciences, University of Leicester, Leicester, UK

4   North Wales Centre for Primary Care Research, Bangor University, Bangor, UK

1 Introduction

The availability of multiple interventions for the same condition or indication is increasingly common [1]. To optimise treatment outcomes and value for money, a sequence of treatments is likely to be used in such contexts. Policy and clinical decisions based on the optimum sequence rather than the effectiveness or cost-effectiveness of discrete treatments are becoming increasingly important [2–5]. This is especially true for chronic diseases, such as depression, diabetes, and cancer [5–7], and some infectious diseases where treatment resistance can limit effectiveness, for example human immunodeficiency virus (HIV) [8]. However, synthesising and interpreting the evidence base to inform such decisions

Key Points for Decision Makers

Treatment sequences, where previous treatment and patient characteristics can affect both the choice and effectiveness of subsequent treatments, are increasingly common in chronic conditions and represent complex treatment pathways. Methods for evidence synthesis that produce the least biased estimates of treatment sequencing effects are required to inform reliable clinical and policy decision making.

Randomised controlled trials (RCTs) of treatment sequences are limited; the use of RCTs of discrete treatments may not provide good evidence on treatment sequencing effects, and observational studies are susceptible to confounding and bias.

The inclusion of discrete treatments used at different points in the treatment pathway may bias a network meta-analysis. Meta-regression needs to account for both previous treatment and duration of disease.

Modelling studies of treatment sequences often apply simplifying assumptions due to the absence of sequencing trials. This can lead to misrepresentation of the true level of uncertainty, potential bias in estimating the effectiveness and cost-effectiveness of treatments, and the wrong decision.

is not straightforward. Treatment sequencing represents a complex intervention pathway where treatment history and patient characteristics may influence both the choice and the effectiveness of subsequent treatments. Treatment history represents multiple factors, including, number and type of previous treatments [9, 10], carry-over effects of prior treatments [11–13], type, level and duration of response to previous treatment [14–16], time on treatment [17], intolerance or toxicity [16, 18], development of disease resistance [19, 20], and burden of preceding treatments that can impact subsequent adherence [7, 21]. Time and disease trajectory are also important factors that can influence the effectiveness of subsequent treatment, the impact of which can be both dependent on and independent of previous treatments [9, 10, 22, 23]. Subsequent treatment choices include dose escalation, add-on therapy, a completely new treatment, or re-use of a previously effective treatment. In some instances, for example relapsing-remitting multiple sclerosis, previous treatments can restrict the choice of allowable follow-on drugs [24].

Randomised controlled trials (RCTs) provide the most robust estimates of treatment effects to inform policy and clinical decision making. However, RCTs of treatment sequences are few in numbers and do not cover the breadth of decision making needed. As the number of available treatments increases, the number of unique sequences will increase geometrically [4, 25], making it impractical and prohibitively costly to evaluate all conceivable sequences in RCTs. The time-varying adaptive nature of many sequences also means that innovative and novel approaches, such as sequential multiple assignment randomised trials (SMARTs), are required for developing the dynamic treatment regimens [26–28]. RCTs of discrete treatments, used at single points in the treatment pathway, provide robust estimates of effectiveness for their specific context, but may not provide representative estimates for these treatments when used in different contexts, such as the later stages of sequences. Participants who enrol into clinical trials and are adherent to discrete treatments may also be quite different from subjects in trials of treatment sequences where alternative, subsequent treatment options are available [7, 29–31]. In sequential treatment studies, participants' decision to end first-line treatment may be influenced by the knowledge there is a second-line treatment readily available [21]. Alternative data sources, which can potentially provide context-specific estimates of treatment effects in different sequences, are longitudinal observational studies, but these are subject to selection bias and confounding.

Evidence synthesis methods that produce the least biased estimates of treatment-sequencing effects are required to inform reliable clinical and policy decision making. Due to the limitations of primary data sources outlined above, this is likely to require advanced meta-analytic techniques [32–36] or mathematical modelling [37]. There is no current guidance for best practice in this context. The Decision Support Unit (DSU), which is commissioned by the National Institute for Health and Care Excellence (NICE) to provide a research resource to support the institute's Technology Appraisal Programme has developed a briefing document on reviewing sequential treatments and downstream costs [38]. This was part of a series of briefing papers and reports developed to inform the 2013 update of the NICE methods guide. The updated methods guide highlighted the fact that some technology appraisals may need to consider the comparison of treatment sequences. However, neither the updated methods guide nor the DSU's briefing document provided guidance on evaluating the clinical effectiveness or modelling treatment sequences. We did not find any other health technology assessment (HTA) guidance that provided information on evaluating treatment sequences. Our paper provides a first step in addressing this limitation.

As a step towards informing best practice, a comprehensive review of reported quantitative evidence synthesis methods was conducted to establish what existing methods are available and outline the assumptions they make and any shortcomings. It is also hoped that this review will draw

attention to this increasingly important area and encourage future methods development.

The review of methods was conducted with the aim of providing guidance for undertaking HTA or similar processes, including comparative effectiveness research and evidence-based guideline development. We did not aim to assess the effectiveness or cost-effectiveness of treatment sequences here, rather the methods used to develop summary treatment effect estimates of whole sequences or discrete treatments conditional on their positioning in the treatment pathway. The review considered methods applied within both clinical and economic evaluation; however, our focus is on the estimation of clinical effectiveness and does not consider the impact of treatment sequencing on the estimation of costs or utility values.

2 Methods

2.1 Literature Search

The intention was to identify the breadth of methods developed for evaluating treatment sequences and not to identify every study that used each method.

The breadth of our review, the recognised challenges of identifying and selecting methodological research using reference databases [39–41], and the fact that the majority of relevant literature would likely be studies reporting applicable methods or methodological developments as part of a wider applied study, rather than being primarily methodological studies [41], meant that a conventional systematic search of reference databases was considered insufficient for the current review. A number of approaches and sources were therefore used to identify relevant methodological studies. The following bibliographic databases were searched from inception to August 2013: MEDLINE, Embase, and the Cochrane Library. The search strategy is provided in Online Resource 1 (see the electronic supplementary material). This was supplemented by hand-searching the following: internet search engines; the websites of specific organisations, including NICE; electronic journals; the agendas of online conference proceedings; the references of existing reviews (listed in Online Resource 1) and relevant papers; known author searches; and forward citation tracking. The reference database searches were not updated, but iterative and purposeful hand searches, including the PubMed related citations function, were continued throughout the review process. An in-depth review was conducted of relevant studies identified during the initial searches. Potential new studies then were then cross-referenced with a list of included studies and recorded methods. More recent studies were only included if they contributed to new methods or knowledge. The searches were deemed to be complete

when further efforts to identify information did not add to the analysis [42] (with the most recent study published in 2016). This is analogous to reaching the point of 'saturation' in qualitative research [42, 43]. The searches have since been supplemented by a recent purposeful and targeted search, which incorporated scanning studies included in a recent systematic review of economic evaluations in rheumatoid arthritis by Ghabri et al. [44].

2.2 Assessing Study Relevance

The review included any disease condition and sequence of any type of treatment. It did not consider decision problems relating to prevention, screening/prognostic, diagnostics, or treatment monitoring. It focused on treatment switching based on a clinical assessment. Studies evaluating the effectiveness of *planned* sequential administration of combined therapy were excluded, as this represented a different type of decision framework.

The review included studies that applied or developed quantitative evidence synthesis methodology as part of secondary research. Studies that used qualitative or narrative evidence synthesis and primary research evaluating treatment sequences were excluded. Any type of meta-analytic technique was considered, incorporating, but not limited to, pairwise meta-analysis, meta-regression, network meta-analysis (NMA), and any meta-analysis based on individual patient data (IPD). Decision-analytic modelling techniques developed to evaluate treatment sequencing, whether conducted as part of an economic evaluation or not, were included. Modelling studies that aimed to evaluate the effectiveness of discrete treatments and incorporated the impact of downstream treatments were only included if they specifically modelled sequencing effects. Studies published in abstract form were excluded, as were economic evaluations based on a single RCT.

3 Results

3.1 Overview of Included Studies

Database searches, after de-duplication, identified 752 references, of which 94 were deemed potentially relevant after screening titles and abstracts. Twenty-six of these could not be further assessed as they were unavailable ($n = 2$), could not be translated ($n = 2$), or were only published as conference abstracts ($n = 22$). A further 28 of those retrieved in full were excluded as they were not relevant (Fig. 1). After collating studies published in more than one publication, the remaining 40 references yielded 36 studies of interest. These were included in the review, along with a further 53 studies identified via internet and hand searches. Recent

supplementary targeted searches identified two studies [45, 46] that contributed a new modelling technique. There were 91 studies in all.

Forty-nine (54%) studies investigated the use of disease-modifying antirheumatic drugs (DMARDs), including biological agents (or biologics), for the treatment of inflammatory arthritis, including rheumatoid arthritis, psoriatic arthritis, and ankylosing spondylitis. Fourteen (15%) related to oncology. The remainder assessed treatments for epilepsy ($n = 4$; 5%), psoriasis ($n = 4$), depression ($n = 3$; 3%), glaucoma ($n = 2$; 2%), schizophrenia ($n = 2$), type 2 diabetes mellitus ($n = 2$), HIV ($n = 2$), neuropathic pain ($n = 1$), postherpetic neuralgia ($n = 1$), sciatica ($n = 1$), fibromyalgia ($n = 1$), chronic hepatitis B infection ($n = 1$), Crohn's disease ($n = 1$), onychomycosis ($n = 1$), and spasticity ($n = 1$). The majority involved sequences of drug treatments, but some also considered other interventions, for example, surgery for sciatica. Only two studies were primarily methodological [14, 47].

Meta-analysis and decision-analytic modelling were reviewed as two distinct categories of quantitative evidence synthesis methods.

3.2 Meta-Analytic Methods

Twenty-three studies were included in the evaluation of meta-analytic approaches [9–11, 16, 23, 47–64]. However, some of these studies were considered relevant in fairly broad terms, such as providing examples of how the limited evidence base precluded the evaluation of treatment sequencing, or representing the use of stratified analysis by line of therapy, which could potentially provide a building block for future methods development. These approaches were initially not considered pertinent to the review but because of the dearth of relevant methods identified, a post hoc decision was made to include them as examples of simplifying methods. This provided a more comprehensive list of the approaches pragmatically used for evaluating treatment sequencing in general, rather than limited to novel methods for developing sequence-specific summary effect estimates. An overview of the studies, including their aims, approaches used, and the data sources, is presented in Table 1.

The evidence to inform treatment sequencing was broadly considered in two ways: a one-step-at-a-time evaluation based on a series of discrete treatments and a comparison of whole sequences. No novel meta-analytic methods (beyond the use of conventional pairwise meta-analysis [32]) were identified for evaluating treatment sequences, and none directly aimed at developing a summary estimate of effect conditional on positioning in the sequence. Most approaches were developed for addressing excessive heterogeneity or specific gaps in the RCT evidence when evaluating discrete treatments at single points in the pathway. For example, in rheumatoid arthritis, RCTs of initial biological therapy investigated the use of these drugs in both early disease, where patients have not previously received any DMARD therapy, and as add-on therapy for established disease in patients with an inadequate response to previous conventional DMARDs, representing a heterogeneous patient population. The first-generation biologics include tumour necrosis factor-α (TNF) inhibitors. Most RCTs of second-line biologics investigated other types of biologics in participants with an inadequate response to previous TNF inhibitors; few RCTs evaluated the sequential use of first-generation TNF inhibitors, whist registry data show that these are often used in practice as second- or subsequent-line therapy [47]. The current meta-analytic approaches, which can potentially be used in a clinical evaluation of a health technology, are summarised below.

3.2.1 Meta-Analysis of Studies Evaluating Whole Sequences

This approach is hampered by the limited number of available RCTs of treatment sequences, which also makes it difficult to establish a closed network for implementing NMA [56]. Observational studies can be used as alternative data sources, but are subject to confounding and bias. The type of observational studies used included the comparison of participants who had received a predefined sequence of two drugs [11], the evaluation of second-line treatment where generic first-line treatment is used [52], and the comparison of the outcomes of first- and second-line treatments [9, 47]. The comparison of treatments used during an earlier versus a later part of the treatment pathway ignores the likely effect of disease trajectory, issues relating to treatment choice, changes in pathophysiology with time, and other confounding factors. The types of bias and limitations of non-randomised studies that are specific to the evaluation treatment sequences, and identified as part of the review, are listed in Box 1.

3.2.2 Subgroup Analyses to Explore the Impact of Treatment History when Evaluating Treatment Sequences in a Piecemeal Fashion

The subgroups can be defined in two ways: by splitting all studies into two or more groups, also referred to as stratified analysis (e.g. early- vs late-stage disease, or failed previous TNF inhibitor 'yes' vs 'no') [63, 64], or by taking partial data from included studies (e.g. participants switching to a second TNF inhibitor due to intolerance, lack of efficacy, or loss of efficacy) [58]. A summary of the methods used is provided in Online Resource 2 (see the electronic supplementary material). Stratified analysis is also applied when
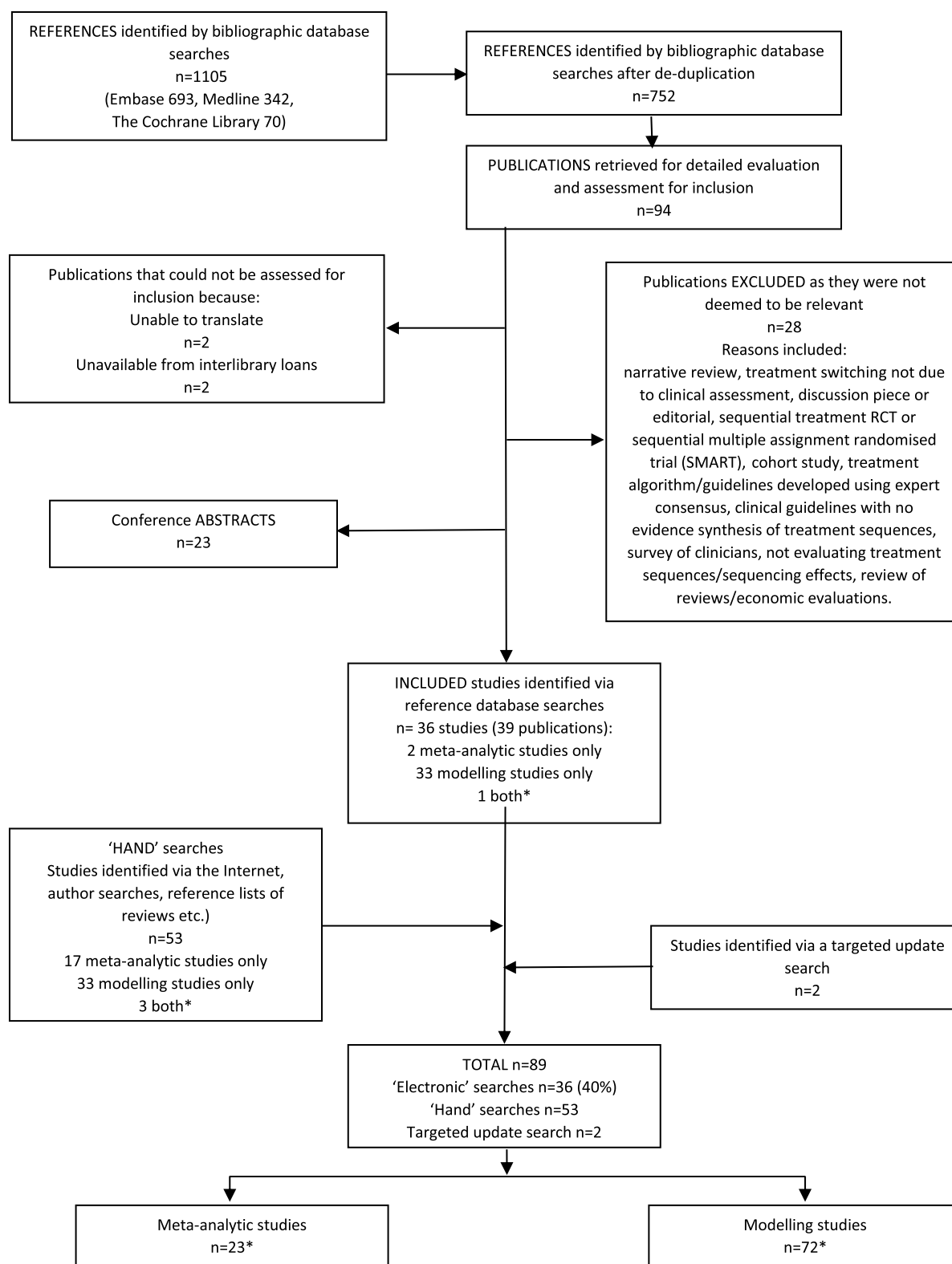
Fig. 1 Flow diagram showing the number of references identified, publications retrieved, and studies included in the methodology review

Table 1 Overview of the meta-analytic approaches used by included studies (studies are ordered according to the methodological approach used)

Study first author, year	Condition	Aim of study	Analysis aimed to evaluate treatment sequences?	Sequencing studies*	Stratified MA (*individual treatment lines*)	Sub-group analyses	Meta-regression	NMA of both 1st- and 2nd-line treatments	Modifying factor	Ranking absolute effects	Available evidence base**
Heng, 2014 [52]	Metastatic renal cell carcinoma	To systematically review published real-world evidence comparing sequential treatments	Yes	X							Observational studies (of sequences)
Stenner, 2012 [11]	Metastatic renal cell carcinoma	To evaluate the optimal sequence for the tyrosine kinase inhibitors sorafenib and sunitinib	Yes	X							Observational studies (of sequences)
Hind, 2008 [53]	Advanced colorectal cancer	Economic evaluation considered planned sequences. Aimed to evaluate the cost-effectiveness of irinotecan, oxaliplatin, and raltitrexed as 1st-line treatments. (*Data on clinical effectiveness limited to planed sequencing trials.*) Clinical evaluation considered treatments used as 1st and 2nd line	Yes	X	X						RCTs of prospective sequences (*n* = 2)

Table 1 (continued)

Study first author, year	Condition	Aim of study	Analysis aimed to evaluate treatment sequences?	Sequencing studies*	Stratified MA (individual treatment lines)	Sub-group analyses	Meta-regression	NMA of both 1st- and 2nd-line treatments	Modifying factor	Ranking absolute effects	Available evidence base**
NICE CG131 [56]	Advanced colorectal cancer	Economic evaluation aimed to assess the effectiveness and cost-effectiveness of chemotherapy sequences	Yes	X	X						1st-line: RCTs; 2nd-line: RCTs of prospective sequences (n = 3, 2 were quasi-sequences)***
Ruhé, 2006 [59]	Major depressive disorder	To systematically review the evidence for switching pharmacotherapy after a first SSRI for major depressive disorder	Yes	X	X						RCTs (one of whole sequences: SMART design) and observational studies; only RCTs pooled due to heterogeneity
Cooper, 2011 [49]	Depression	To systematically review studies of the management of treatment-refractory depression in older people, covering pharmacological, physical, and psychological interventions (Search strategy included treatment sequencing)	Yes	X	X						RCTs and uncontrolled open-label trials (one non-RCT sequences study); pooled breaking randomisation
Lloyd, 2010 [9]	Rheumatoid arthritis	To evaluate the effectiveness of TNF inhibitors when used sequentially	Yes	X		X	X				Observational studies (uncontrolled and comparative studies of 1st- vs 2nd-line studies)

Table 1 (continued)

Study first author, year	Condition	Aim of study	Analysis aimed to evaluate treatment sequences?	Sequencing studies*	Stratified MA (individual treatment lines)	Sub-group analyses	Meta-regression	NMA of both 1st- and 2nd-line treatments	Modifying factor	Ranking absolute effects	Available evidence base**
Rendas-Baum, 2011 [58]	Rheumatoid arthritis	To evaluate the relationship between the clinical response to biologics and the number of previous treatments with TNF inhibitors	Yes			X					RCTs and observational studies; pooled breaking randomisation
Suarez-Almazor, 2007 [64]	Rheumatoid arthritis	To review the evidence on the TNF inhibitors, INF and ETA, regarding the timing of therapeutic introduction, dose escalation, and switching	Yes		X	X					1st-line: RCTs; 2nd-line: observational studies only (not pooled)
Schoels, 2012 [62]	Rheumatoid arthritis	To compare efficacy and safety of biologics after inadequate response to TNF inhibitors (Provided an example of the limitations of examining differences in treatment effects according to the number of previous biologics used.)	No			X					RCTs

Table 1 (continued)

Study first author, year	Condition	Aim of study	Analysis aimed to evaluate treatment sequences?	Sequencing studies*	Stratified MA (individual treatment lines)	Sub-group analyses	Meta-regression	NMA of both 1st- and 2nd-line treatments	Modifying factor	Ranking absolute effects	Available evidence base**
Singh, 2009 [63]	Rheumatoid arthritis	To compare efficacy and safety of biologics Included planned subgroup analyses for TNF failure vs none; and biologic failure vs conventional DMARD failure vs none	No			X					RCTs
Salliot, 2011 [60]	Rheumatoid arthritis	To compare efficacy of biologics in two clinical situations: (1) active disease despite MTX; (2) after inadequate response to TNF inhibitor (*Provides an example of evaluating effect of previous treatment in subgroup analysis*)	No		X	X					RCTs

Table 1 (continued)

Study first author, year	Condition	Aim of study	Analysis aimed to evaluate treatment sequences?	Sequencing studies*	Stratified MA (individual treatment lines)	Subgroup analyses	Meta-regression	NMA of both 1st- and 2nd-line treatments	Modifying factor	Ranking absolute effects	Available evidence base**
Nixon, 2007 [57]	Rheumatoid arthritis	To compare the efficacy of four biologics, three of which were TNF inhibitors (*Provides an example of using MR to account for disease duration in NMA; first study to combine MR and NMA*)	No				X				RCTs
Schmitz, 2012 [61]	Rheumatoid arthritis	To compare efficacy of TNF inhibitors in patients with inadequate response to MTX (*Provides an example of the challenges of including previous treatments as a covariate in MR due the the poor reporting of primary studies; represents NMA combined with multivariate MR*)	No				X				RCTs

Table 1 (continued)

Study first author, year	Condition	Aim of study	Analysis aimed to evaluate treatment sequences?	Sequencing studies*	Stratified MA (individual treatment lines)	Subgroup analyses	Meta-regression	NMA of both 1st- and 2nd-line treatments	Modifying factor	Ranking absolute effects	Available evidence base**
Christensen, 2015 [10]	Rheumatoid arthritis	To determine if variations in trial eligibility criteria and patient baseline characteristics could be considered effect modifiers of the treatment response when testing targeted therapies (biological and targeted synthetic DMARDs) (*Factors considered included previous DMARD and disease duration*)	No				X				RCT

Table 1 (continued)

Study first author, year	Condition	Aim of study	Analysis aimed to evaluate treatment sequences?	Sequencing studies*	Stratified MA (individual treatment lines)	Subgroup analyses	Meta-regression	NMA of both 1st- and 2nd-line treatments	Modifying factor	Ranking absolute effects	Available evidence base**
Kanters, 2014 [54]	Rheumatoid arthritis	To explore which clinical factors and patient characteristics are associated with the magnitude of comparative efficacy between biologics vs MTX patients with inadequate response to MTX (*Demonstrates the challenges of including previous treatments as a covariate in MR due to the poor reporting of primary studies*)	No				X				RCTs
Anderson, 2000 [23]	Rheumatoid arthritis	To identify factors predicting response to 2nd-line treatment, with conventional DMARDs or devices (*Factors considered included previous DMARD and disease duration*)	No				X				RCTs (individual patient-level data)

Table 1 (continued)

Study first author, year	Condition	Aim of study	Analysis aimed to evaluate treatment sequences?	Sequenc-ing stud-ies*	Stratified MA (*individual treat-ment lines*)	Sub-group analyses	Meta-regres-sion	NMA of both 1st- and 2nd-line treatments	Modi-fying factor	Ranking absolute effects	Available evidence base**
Mandema, 2011 [55]	Rheumatoid arthritis	To compare the dose–response relationship for the efficiency of biologics. Two of the objectives included the following: Are TNF inhibitors different in patients with an inadequate response to MTX compared to those who are MTX-naïve? Are TNF inhibitors more efficacious than MTX in MTX-naïve patients?	No		X		X				RCTs
Grothey, 2004 [51]	Advanced colo-rectal cancer	To evaluate the importance of the availability of all three active cytotoxic agents, FU-LV, irinotecan, and oxaliplatin, on overall survival. Stand-ard 1st-line therapies were FU-LV plus irinotecan or oxaliplatin	Partial				X				RCTs (one of sequences)****

Table 1 (continued)

Study first author, year	Condition	Aim of study	Analysis aimed to evaluate treatment sequences?	Sequencing studies*	Stratified MA (individual treatment lines)	Subgroup analyses	Meta-regression	NMA of both 1st- and 2nd-line treatments	Modifying factor	Ranking absolute effects	Available evidence base**
Abrams, 2016 (IMI GetReal Project case study: https://www.imi-getreal.eu/) [47]	Rheumatoid arthritis	To explore how real-world data, from patient registries, can be used to help demonstrate the relative effectiveness of new medicines. Addresses two key issues: How to connect disconnected networks of evidence to conduct NMA; How to optimise an evidence base using 1st-line evidence to inform 2nd-line effectiveness estimates	Partial					X			RCTs and patient registries (individual patient-level data); registry data used to develop comparative studies of 1st vs 2nd lines
Rodgers, 2011 [16]	Psoriatic arthritis	To determine the clinical effectiveness, safety, and cost-effectiveness of TNF inhibitors in the treatment (*Sensitivity analysis for economic model incorporated sequential TNF inhibitors*)	Yes						X		Observational studies

Table 1 (continued)

Study first author, year	Condition	Aim of study	Analysis aimed to evaluate treatment sequences?	Sequencing studies*	Stratified MA (individual treatment lines)	Subgroup analyses	Meta-regression	NMA of both 1st- and 2nd-line treatments	Modifying factor	Ranking absolute effects	Available evidence base**
Connock, 2006 [48]	Epilepsy	To examine the clinical effectiveness and cost-effectiveness of newer antiepileptic drugs for epilepsy in children. (Economic model included treatment sequencing)	Yes						X		RCTs
Finnerup, 2005 [50]	Neuropathic pain	To develop up-to-date calculation of NNT and NNH as the basis of a proposal for an evidence-based treatment algorithm (sequence of treatments)	Partial							X	RCTs

CG clinical guidelines, *DMARD* disease-modifying antirheumatic drug, *FU-LV* fluorouracil-leucovorin, *IMI* Innovative Medicines Initiative, *MA* meta-analysis, *MTX* methotrexate, *NICE* National Institute for Health and Care Excellence, *NMA* network meta-analysis, *NNH* numbers needed to harm, *NNT* numbers needed to treat, *RCT* randomised controlled trial, *SMART* sequential multiple assignment randomised trial, *TNF* tumour necrosis factor

* Whole sequences or comparing treatment lines

** Unless otherwise stated, RCTs relate to the evaluation of discrete treatments; 'placebo RCTs' included a placebo control, whilst 'RCTs' included either an active or placebo control

*** Quasi-sequencing trials: RCTs of 1st-line treatment with subsequent treatment predefined in protocol, or high proportion of patients went on to receive the same 2nd-line treatment

**** Included published RCTs that reported the number of patients receiving 2nd-line therapies made by the authors of the trials

conducting separate meta-analysis for each line of therapy (e.g. first- or second-line biologics) [62, 64] or for different patient populations (e.g. participants with no previous history of biologic therapy or participants with an inadequate response to previous TNF inhibitors) [10, 60]. The main limitation of using subgroup analyses is that it only allows for the comparison of two subgroups at a time, with or without one covariate. All other covariates are pooled, and each analysis is confounded by other variables [65].

3.2.3 Meta-Regression to Adjust for the Previous Treatment

This approach was not generally used for the sole purpose of evaluating treatment sequences, but was used to account for the heterogeneity within the meta-analysis or NMA. A summary of the methods used is provided in Online Resource 2. The covariate representing previous treatment was often dropped from the final analysis due to non-significant findings [54, 61], possibly due to lack of power, as previous treatment was generally poorly reported in primary studies [10, 54]. However, lack of variability between studies can also contribute to non-significant findings, especially when the meta-analysis is used to compare treatments applied at a single point in the pathway, or where the ordering of treatments is much the same in a given disease. To avoid problems with insufficient power, a limited number of covariates are incorporated in the meta-regression. This frequently included disease duration. For example, a study, which combined the use of NMA and meta-regression to account for the significant heterogeneity between studies of biologics for rheumatoid arthritis, included only two study-level covariates in the meta-regression, disease duration and a measure of baseline disability [57]. The analysis included RCTs of participants who were DMARD naive and RCTs of participants with an inadequate response to these drugs lumped together. Disease duration could potentially be considered as a proxy for previous treatment use, as the likelihood of failing prior treatments will increase with increasing duration. However, there is also justification for including treatment history as a covariate, especially when pooling (lumping together) studies across different treatment lines [10, 23]. The inclusion of both covariates could help to disentangle whether long standing disease per se is associated with a poor response to treatment, or whether failure on previous treatments predicts response to subsequent treatment [22]. The use of IPD is likely to enhance the application of this approach [10], but studies that used such data were still hampered by the poor reporting of previous treatment [23].

A further limitation of conducting an NMA of all discrete treatments irrespective of where they are used in the pathway is that previous treatment(s) can both have an impact on treatment effect, acting as an effect modifier, resulting in heterogeneity, and be associated with the type of treatment comparison, acting as a confounding factor and lead to inconsistency in the network. For example, in an NMA of sciatica treatments, non-invasive treatments were more likely to be used as initial treatments and invasive treatments were used after the failure of other treatments in patients with a more long-standing and less responsive condition [66].

3.2.4 Network Meta-Analysis Incorporating Multiple Treatment Lines, for Example, First- and Second-Line Treatments, as Separate Treatment Nodes

This approach was not developed for evaluating treatment sequences as such, but rather to evaluate methods for incorporating real-world data in evidence synthesis of second-line treatment. In particular, the approach sought to optimise an evidence base using first-line evidence to inform second-line effectiveness estimates. The methods were applied as part of the GetReal project case study of biologics in rheumatoid arthritis [47]. The authors had access to IPD from two national registries and five RCTs (two investigated second-line treatment). A series of Bayesian univariate and bivariate NMA was conducted that incorporated both treatment lines. The data from RCTs provided separate networks of evidence for first- and second-line biologics. No RCT reported on both treatment lines; thus the exchangeability assumption was needed to connect the two networks by assuming all treatment effects have a common distribution. The univariate analysis utilised the registry data as data, whereas the bivariate analyses used the registry data to inform the prior distribution for the correlation parameter between first- and second-line treatments. In the univariate analysis, relative effect estimates for first- versus second-line treatment were obtained from the registry, allowing the two networks to be connected and for treatment comparisons (e.g. drug A in first line vs drug A in second line) to be obtained. The use of multivariate analysis allows separate outcomes to be modelled simultaneously, using the correlation to borrow information across multiple outcomes or time points. Here, the treatment effect for first-line treatment was modelled as outcome 1 and second-line treatment as outcome 2, and the correlation was that of between treatment lines. The initial bivariate NMA was conducted using RCTs of first- and second-line treatments. The correlation estimate was obtained by conducting standard pairwise meta-analysis, based on registry data split into first- and second-line response, and monitoring the correlation. In a second bivariate analysis, the registry data were used as part of the NMA by being split into multiple pairwise studies. This allowed for modelling between-study correlation between the lines of therapy. A third analysis used data from the registries, reporting treatment effect estimates on both lines, which allowed for relaxing the exchangeability assumption on the average level. The biggest challenge here was developing an estimate of

correlation between first- and second-line treatments to conduct the analysis. The assumptions of consistency and similarity, across the pairwise contrasts, within the NMA may also be difficult to justify, as discussed above in the NMA of sciatica treatments example. The limitations of relying on observational studies comparing first- and second-line treatment are discussed in Sect. 3.2.1 and Box 1.

3.2.5 Developing a Specific Multiplication Factor that Can be Applied to the Summary Effect of a Treatment Used as First Line in Order to Represent Its Use at a Later Point in the Pathway

This approach is not a meta-analytic method as such, but was used to adapt the findings of meta-analysis of discrete (first-line) treatments to represent sequencing effects. The optimal approach for developing a multiplication factor is yet to be established. Current methods incorporate two approaches [16, 48]. One study, investigating TNF inhibitors for psoriatic arthritis, obtained modifying factors from an observational study comparing the class of drugs used as first-line and subsequent treatment for rheumatoid arthritis. A different multiplication factor was developed, depending on whether the initial TNF inhibitor was discontinued due to inefficacy or adverse effects [16]. A second study developed a reduction factor based on the data available for one antiepileptic drug for which there was an RCT of its use at two different time points, first-line monotherapy and later as an add-on therapy [48]. Modification factors were primarily used by modelling studies, with most not reporting the methods used for developing them [18, 67–72]. Most used estimates based on available evidence, mainly an observational or previous modelling study, the choice of which was frequently not justified. The reduction factor used in the most recent (2020) economic evaluation [45] was obtained from a pragmatic RCT of non-TNF-targeted biologic versus a second TNF inhibitor to treat rheumatoid arthritis in patients with insufficient response to their first anti-TNF-inhibitor (Gottenberg et al. [73]).

3.3 Decision-Analytic Modelling

3.3.1 Decision Modelling Methods

Seventy-two modelling studies were reviewed and fifty-two distinct models identified [14–18, 45, 46, 48, 53, 56, 67–72, 74, 74–101, 101–127]. An overview of the included modelling studies is provided in Online Resource 3 (see the electronic supplementary material). Most modelling studies were conducted as part of an economic evaluation. A wide range of modelling techniques were used to address a broad spectrum of treatment-sequencing decision problems (Box 2), which included identifying the optimum sequence; adding a

new drug to an established sequence; comparing 'step-up' or 'step-down' treatment approaches; comparing different treatments used at the same point within a sequence; evaluating a drug used at different points within a sequence; and comparing predefined sequences. The sequence of treatments being modelled ranged from a fixed sequence of a limited number of treatment lines to variable treatment algorithms where patient history and characteristics dictate the choice of subsequent treatments.

Two published taxonomies developed for categorising different modelling techniques according to their key features [128, 129], along with other guides and algorithms that have been developed to aid the selection of an appropriate modelling technique (or structures) for economic evaluation in general [97, 128–140], were used to categorise the included studies and inform the data extraction. The advantages and disadvantages of each modelling approach were assessed as part of the review. The choice of an appropriate modelling approach depends on the complexity of the underlying decision problem, the extent of the treatment sequences being investigated, and the disease condition. Table 2 provides an abbreviated summary of the overall findings of the review of modelling studies, including how treatment sequences were conceptualised within different modelling approaches (column 2); and the type of the additional attributes in the decision problem (beyond the sequencing of individual treatments) and disease condition that were captured by the included models (column 3). A more detailed summary of the methods and findings of the review of modelling studies is provided in Online Resource 4 (see the electronic supplementary material). The modelling techniques used included deterministic decision tree, stochastic decision tree, Markov cohort model, partitioned survival cohort model, semi-Markov cohort model, individual-patient simulation state transition models, discrete event simulation, discretely integrated condition event (DICE) simulation, non-terminating population-based simulation, terminating population-based simulation, and dynamic Markov cohort model. No study compared any of these alternative approaches for evaluating treatment sequences to assess, for example, how sensitive results were to the type of model used. A number of studies did report choosing a discrete event simulation over a state transition model due to the improved computational efficiency [48, 68, 104, 122]. The level of complexity in the decision problem accounted for in the models varied quite considerably, even when evaluating similar treatment sequences within the same disease condition. The decision problem was also simplified by modelling a limited number of treatment lines, streamlining the disease process, and using a short time horizon. For example, some studies used a 2–5 year time frame, rather than a lifetime horizon, for modelling treatment sequences for rheumatoid arthritis, because

a longer time horizon implied too many assumptions [71, 78, 79, 84, 112, 113, 124].

3.3.2 Simplifying Assumptions Regarding Sequences of Treatment

Treatment sequences were often modelled as a series of discrete treatments, each requiring a summary effect estimate conditional on positioning in the treatment pathway. The scarcity of data to inform such estimates meant that simplifying assumptions were often applied to the available data on discrete treatments used at a single point in the pathway. A range of simplifying assumptions used to represent treatment-sequencing effect estimates was identified, which were used to develop a novel taxonomy of all possible assumptions (Table 3). The most common assumptions were that treatment effect is independent of positioning in the sequence, or that treatment effect is dependent on the number of previous treatments (treatment line), but independent of the type of treatments used (Table 4). These assumptions were frequently not validated; nor was their impact on the overall results assessed. Forty-nine studies (72%) assumed that the effect of either all or some of the treatments used after the first treatment modelled (or decision point) were independent of treatment sequence. Only six studies (9%) evaluated the impact of applying this assumption in sensitivity analyses, by reducing the effect of treatments used later in the sequence using a factor based on evidence [67, 69], an arbitrary amount [15, 93, 110], or expert consensus [14]. The assumption that treatment effect is dependent on line of therapy was often used in conjunction with the assumption of treatment independence, applied to treatments adopted later in the sequence.

The available evidence to inform treatment-sequencing effects impacts the type of assumptions required. The review focused on modelling studies that evaluated treatment sequences, but economic evaluations often focus on the comparison of discrete treatments and model downstream costs of subsequent treatments. The findings demonstrated that priority was often given to matching the evidence for the decision point, for example, comparing first-line biologics, rather than considering treatment sequences as a whole. Economic evaluations undertaken by, or on behalf of, manufacturers of health technology tended to focus on a specific decision point reflecting treatments used in pivotal RCTs matching the licence indication, for example, comparing a TNF inhibitor to a conventional DMARD [74, 80, 101], or a non-TNF-inhibitor biologic to a TNF inhibitor [101]. The data sources used alongside the simplifying assumptions for treatments used beyond the decision point varied, even when considering the same decision problem and addressing the same evidence gap. For example, data sources used to inform sequential TNF inhibitors included the following:

RCTs of TNF inhibitors used as first-line biologics [45, 67, 72, 83, 87, 89, 92, 96, 101, 109], a national patient registry [81, 101, 104, 115, 122], a large, uncontrolled, open-label study of a specific TNF inhibitor in patients who had previously discontinued TNF inhibitors [78, 79, 84, 107, 112, 113], and an RCT of a non-TNF-inhibitor biologic in participants with an inadequate response to TNF inhibitors. The effects of treatments administered later in the treatment pathway were also handled in different ways. For example, in a technology appraisal of TNF inhibitors for rheumatoid arthritis [83], the initial treatment response for each subsequent conventional DMARD was explicitly modelled, whilst in another technology appraisal of TNF inhibitors for psoriatic arthritis [16], the economic model assumed that patients experienced a steady long-term deterioration after the failure of the TNF inhibitor. Therefore, fluctuations caused by response to subsequent conventional DMARDs, which were considered to be administered as part of palliative care, were ignored. The uncertainty in the quality of the alternative evidence to inform sequencing effects was not investigated in depth.

Decision models that start at the point of diagnoses are more likely to reflect the complete sequence of treatments used in chronic conditions, for example, some studies of biologics in rheumatoid arthritis developed the decision population within the actual model, with patients entering the model being newly diagnosed with early disease [67, 75, 83, 85, 99, 122]. However, the likelihood that there is no matching evidence is increased, and more assumptions are required. Another approach is to model the initial treatment used prior to the decision point (e.g. when comparing second-line biologics), and apply the assumption that the entire patient population on entering the model have an inadequate response to the first modelled treatment (e.g. first-line biologic). This approach was used in the Advanced Simulation Model, to allow the initial treatments to be costed appropriately, reflecting treatment sequences used in practice [71, 78, 79, 84, 112, 113]. However, the evidence used to inform the treatment effects of the second TNF inhibitor did not match the prior TNF inhibitor failed (first treatment modelled). The third and most common approach was to include a patient population entering the model that reflected the decision problem in terms of the number of previous treatments used, for example, patients receiving their first biologic therapy. Modelling studies that only consider the impact of subsequent treatments when, for example, comparing first-line biologics [71, 72, 74, 75, 83, 89, 101, 120, 123, 127] are generally based on the assumption that the sequences being compared are starting from a level playing field. The potential impact of this is not generally considered within the sensitivity analysis, as it is not part of the cost-effectiveness estimates.

A frequent problem when evaluating the introduction of a new treatment to an established sequence is the lack of data to inform the 'displaced effect'. For example, when adding a new drug (e.g. non-TNF-biologic agent) to an established sequence (e.g. starting with a TNF inhibitor), the existing drug is displaced lower down the sequence (Box 2), and is generally modelled as both the comparator (e.g. first-line) treatment in the baseline sequence and the subsequent (second-line) treatment, after the new drug, in the intervention sequence. The same treatment effect is generally applied to the existing drug, irrespective of whether it is used early or later in the sequence (and disease trajectory), with no RCT data available on its effect in patients with an inadequate response to the new drug.

4 Discussion

4.1 Summary of the Findings

The review identified a range of quantitative evidence synthesis methods used for evaluating the effectiveness of alternative treatment sequences. The findings demonstrated the following:

i   Reviewing the evidence on treatment sequencing is neither trivial nor straightforward.

ii In most cases, treatment sequences represent complex, multifaceted, dynamic intervention pathways, which will require advanced methods of quantitative evidence synthesis, especially if evaluated using a 'one-step-at-a-time' approach.

iii Prospective sequencing trials are few in numbers and do not cover the breadth of decision making needed. The evidence synthesis would likely need to consider the inclusion of diverse study designs, including non-randomised studies.

iv The problem has largely not been addressed using evidence synthesis methodology for clinical effectiveness, but is usually dealt with at the decision modelling stage.

v. There is no single best way to evaluate treatment sequences; rather there is a range of approaches and, as yet, no generalised methodology that encompasses the different assumptions used.

vi Each approach has advantages and disadvantages and is influenced by the evidence available and decision problem.

vii When using a one-step-at-a-time approach, previous treatment is an important effect modifier, and subsequent treatments can confound long-term outcomes, such as survival.

viii The reason for discontinuing treatment (lack of effect, loss of effect, or intolerance) has a differential effect on

the effectiveness (and choice) of subsequent treatment, and is poorly reported in primary studies.

ix The extent and type of sequences being evaluated tended to reflect the available research evidence, rather than clinical practice.

4.2 Comparison with Existing Reviews

We identified three existing reviews of methods for evaluating treatment sequences. This included two systematic reviews of economic evaluations [4, 141] and one review of published UK NICE technology appraisals [3]. Mauskopf et al. analysed treatment-sequencing assumptions after failure of the first biologic in cost-effectiveness models of psoriasis, and compared the modelled sequences with the most recent treatment guidelines [141]. They concluded that models of first-line biologics either do not include subsequent treatments or include only some of the regimes recommended in current guidelines, and that cost-effectiveness results may be sensitive to the assumptions about treatment sequencing, and choice and efficacy of subsequent treatment sequencing regimens. Tosh et al. assessed and critiqued how sequential DMARDs for rheumatoid arthritis have been modelled in economic evaluations [4]. They found that reporting of the methods and evidence used to assess the effect of downstream treatments was generally poor; when lifelong models and treatment sequences were considered, evidence gaps were identified. They concluded that methods were not applied consistently, leading to varied estimates of cost-effectiveness, and that treatment sequences were not fully considered and modelled, potentially resulting in inaccurate estimates of cost-effectiveness. Zheng et al. investigated approaches used to model treatment sequences in NICE appraisals to provide practical guidance on conceptualising whether and how to model sequences in health economic models [3]. They concluded that the biggest challenge is the scarcity of clinical data that capture the long-term impacts of sequences on efficacy and safety. Three commonly used assumptions to bridge the evidence gap were identified, but each had its own limitations. These included the assumption that the efficacy of a treatment stayed unchanged regardless of line of therapy, the use of data from trials in different lines of therapy to directly model a treatment sequence, and the use of retrospective studies of clinical registries or databases. The findings of these reviews were consistent with ours, though their scope was more limited in that they focused on either a single condition or UK NICE appraisals.

4.3 Strengths and Limitations of the Review

This is the first review of methods to investigate the evaluation of treatment sequencing across all clinical scenarios,

---

Box 1 Potential bias or limitation in non-randomised, real-world observational studies that are specific to the evaluation of treatment sequences

---

i. **Selection (allocation) bias** results in systematic differences in prognostic factors between individuals in treatment and control groups, e.g. a cohort of patients receiving their first tumour necrosis factor (TNF)-inhibitor compared with a cohort receiving a second or subsequent TNF-inhibitor. Patients in the second group are likely to have worse prognoses and show limited responses to all treatments [9, 175]. Adjustment for both baseline and post-baseline prognostic factors is necessary, to ensure the comparability of treatment groups [166].

ii. **Channelling bias** favours patients with more severe disease. New treatments create expectations of improved effectiveness and tolerability; early, post-marketing users are likely to be those who experienced little or no benefit from existing drugs and may therefore respond to the new drug in a way that is not representative of the eventual user population [176].

iii. **Regression to the mean** occurs because patients tend to be treated with a second or subsequent treatment at the height of their disease activity, where there is a greater than 50–50 likelihood that the condition will start improving after the intervention purely by chance [22].

iv. **Confounding by disease duration** occurs in conditions such as sciatica and rheumatoid arthritis, where the longer the disease duration, the less likely that patients will respond to any treatment [14]. Treatment history can be both correlated with disease duration [10] and act as independent effect modifier [10].

v. **Enrichment of successive treatment use with refractory patients** A small proportion of patients have refractory disease that will not respond to any treatment [14]. Populations receiving second-line or subsequent treatments are more likely to be enriched with such patients. This is related to class effect bias (vii). Patients who fail initial treatment due to a tolerability or safety issues are likely to have the same problem with any alternative drug from the same class, increasing the risk for developing an adverse event in patients who switched due to an adverse event [16].

vi. **Immortal time bias** occurs in studies that limit inclusion to patients who are receiving a specific line of treatment (e,g, third-line chemotherapy) [52] or have completed a predefined sequence, and overlook patients who are continuing the initial treatment, or lost to follow-up after first-line treatment due to lack of efficacy, clinical deterioration, death or drug acceptability issues [52]. It is particularly relevant for treatments of advanced cancer where a large proportion of patients may not complete the sequence, or receive multiple treatment lines [52].

vii. **Class effect bias**, which is the possibility that the comparison between drug classes may be confounded by differences in the type of patients treated with each class [52].

viii. **Aggregate data collection** is a limitation of real-world observational studies that do not report individual treatment or drug-level data. Any subsequent evidence synthesis has to be based on pooled data across treatments at class level, even when there is evidence that individual drug effects can vary within a class [11]. Class level treatment effects are often reported even when access to individual patient data is available [82].

ix. **Missing or inaccurate data** may be obtained from real-world practice. Patient registers and administrative databases are rarely set up for evaluating treatment-sequencing, and may not involve a high level of rigour in recording events [52].

x. **Variability in how the same outcome measure is conceived across different studies** is a particular issue in oncology when using progression free survival (PFS) to evaluate the impact of a sequence of treatments (e.g. using the sum of the progression free survival period for each treatment line). PFS is a composite endpoint, which may or may not incorporate a treatment free period before the next treatment resulting in a differential impact on the results; this needs to be accounted for in any pooled analysis [11]. Importantly, the use of PFS associated with each successive treatment line to inform treatment-sequencing assumes that all treatment effect from each treatment line stops on progression [11]. Alternative endpoints that have been proposed for evaluating a fixed sequence of treatments [183] include: Duration of disease control (DDC) and Time to failure of strategy (TFS).

NB The type of biases listed here may not be mutually exclusive and the descriptors may not be consistently used in the published literature, for example the phenomenon described as `regression to the mean' can also be representative of both a class effect and a channelling bias, favouring patients with more severe disease [180].

---

and to include both meta-analytic techniques and decision-analytic modelling. It represents an extensive in-depth review of current methods used to evaluate the clinical effectiveness of treatment sequences, representing a broad and disparate area of research.

A potential limitation of our review is that the reference database searches were not updated. However, targeted hand searches were continued during the review process and studies published beyond 2013 have been included. Nevertheless, more recent studies were only included if they contributed new findings, and the searches stopped when no new information was being found. This means that the review could have potentially missed new methods developed in the last few years. Updated targeted hand searches identified a new modelling technique (DICE) that was not previously included in our review. This has since been included. However, the methods used to conceptualise treatment sequences and the level of reality captured in the DICE model did not change the findings and recommendation of our review. The methods used to develop treatments sequencing effect estimates and the accompanying simplifying assumptions made within the new studies [45, 46] were also the same as those included in our review. The assessment of recent studies included in a new systematic review of economic evaluation of sequences of biological treatments for patients with rheumatoid arthritis, published in 2020, did not identify any

studies reporting methods or simplifying assumptions not already incorporated in our review [44].

4.4 Recommendations for Practice

4.4.1 Primary Study Design

The available evidence base for evaluating new treatments is often driven by the requirements for regulatory approval, and thus focuses on discrete treatments used at a defined point in the pathway [142, 143]. The lack of data on the effectiveness of these treatments when used at another point in the pathway is a barrier to making policy decisions about the optimal positioning of new treatments or treatment sequences. The GetReal project (Sect. 3.2.4) included a stakeholder engagement workshop to solicit views on the usefulness and acceptability of their analytic approach [144]. Interestingly, the regulators considered it to have limited usefulness because the evidence requirements for marketing authorisation in rheumatoid arthritis is line specific, whilst the pharma research and development representatives considered it useful in principal, to better understand the gaps in the evidence across lines of therapy and aid the design of future clinical trials [144]. The focus of primary research on discrete treatments is unlikely to change unless the regulatory authorities specify the importance of treatment sequencing or optimal positioning of new treatments. The reimbursement agencies and HTA bodies should also make recommendations on the nature of the clinical evidence required to inform treatment sequences [145, 146].

4.4.2 Health Technology Assessment

It is important to identify the relevance of, or the need to consider, treatment sequencing early on in the technology assessment process, and incorporate both the clinical and economic evaluation. Treatment sequencing was often considered as part of the economic evaluation only, and not considered in the clinical evaluation [17, 67, 83, 85, 95, 99, 106, 116, 126, 147]. A previous review of NICE technology appraisals also identified a lack of integration or direct use of the systematic review to inform the economic evaluation, and the need to consider the data requirement of the economic model at an early stage [148].

The development of an initial analytic or conceptual framework [40, 149] provides an essential tool for the planning and evaluation of treatment sequences. It can be used to consolidate the requirements of the clinical and economic evaluation; assist in communication within the research team and with a range of stakeholders; 'think through' the multiple components of the treatment pathways and disease-specific events in context; enhance the transparency of underlying assumptions; and inform choices about the level of structural complexity required by the model [40, 139, 150–153]. For some chronic diseases, it may be useful to create a disease-specific conceptual framework that can serve as a foundation for developing future HTAs and economic models of current and novel treatments [154], thus potentially allowing for greater stakeholder feedback and future improvement. There is also a need to depict treatment sequences as a tree, rather than a linear sequence of treatments, thus accounting for the complex and dynamic intervention pathways that they represent. Although methods were developed that accounted for the fact that the reason for treatment discontinuation (e.g. loss of effectiveness, adverse events, non-adherence) might determine the average effectiveness for the next line of therapy, the reality is that this may also affect the choice of therapy for the next line. A tree structure is adopted in the SMART design, which is a multistage trial designed to develop and compare treatment pathways that are adapted over time based on individual's response and/or adverse effects [28].

The time and resource constraints of HTA, accompanied by limited evidence, may render an extensive model unrealistic. It may therefore be tempting to simplify the treatment-sequencing decision problem. However, a model based on an oversimplification of the decision problem and clinical practice is also unlikely to be useful for decision makers. An alternative approach would be to develop a model that is designed to address any/multiple decision problems, rather than a single use model. This may be relevant, not only for chronic disease, but also in the introduction of new treatments in a rapidly changing clinical field, such as oncology [5]. The likelihood that the available data to inform sequencing effects may improve over time also supports developing a model that is easily updated. This is consistent with recent calls for the use of disease-specific reference models [155], pre-verified modules [156], and open-source models [157] to improve the accuracy of economic evaluations. Our review identified some good examples where a model was further developed over time to address multiple reimbursement decisions (e.g. Birmingham Rheumatoid Arthritis Model [BRAM] [75, 158], Tran-Duy model [68, 122], Sheffield rheumatoid arthritis models [159], and the Advanced Simulation Model [78]). However, each was developed by the same research group. An important challenge here is the need to make sufficient detail on the original model openly available.

A mathematical challenge for comparing multiple permutations of sequences is to determine the proper starting point of the model. This is also relevant when using a model designed for multiple uses, which may start at the point of diagnosis [75], a key point in the treatment pathway (e.g. initiating DMARD therapy [122] or biologic therapy [78]), or the point at which the decision is made (Sect. 3.3.2). All evaluations should start at the point of divergence (i.e. the

---

Box 2 Illustration of the different types of treatment-sequencing decision problems

---

As part of the review of modelling studies, a coding scheme was developed for categorising modelling studies according to the type of decision problem relating to treatment sequences that was evaluated. The codes used are illustrated below. Some studies include more than one decision problem type.

a). **'Optimum sequences'**

Identifying the best sequence out of all conceivable sequences (as opposed to comparing predefined sequences, thus selecting a manageable number of sequences for comparison in advance)

b) **'Predefined sequences'**

A - B - C

B - A - C

X - Y - Z

Comparison of pre-specified sequences; also incorporates the following:

c) **'Disease approach'**

A - B

B - A

or

X - A - B

A - B - **X**

Comparison of 'step-up' vs 'step-down' approaches, or the use of new drugs first vs starting with older, established drugs.

d) **'Single point'**

A - B - C - D

A - B - **X** - D

Comparison or decision point = C vs X. Treatment C is replaced by X in the second sequence

e) **'Different points'**

X - B - C - D

A - **X** - C - D

A - B - **X** - D

A - B - C - **X**

Comparison of X used at different points in the sequence

f) **'Adding' a new treatment to an established sequence**

A - B - C - D

A - B - **X** - C - D

Comparison or decision point = C vs X. Treatment C is displaced by X in the second sequence

---

point at which a decision might be made) [75]. Models used for comparing multiple permutations of sequences often include the same first one, two, or three lines of treatment. This will essentially 'dilute' the true incremental effects (and costs) of treatment since some patients will have died (and left the model) before the point of divergence. Thus, when calculating the incremental outcomes per patient, the denominator will be greater than should have been used, meaning that the incremental results will underestimate the true effects.

A number of studies developed a model based on an existing approach. Existing modelling approaches could also, potentially, be adapted for use in a different disease condition. However, when using an existing model, it is important to consider what underlying assumptions regarding treatment sequences were applied. For example, the York psoriasis model [126], which has subsequently been used by multiple studies evaluating treatment sequences in psoriasis [141, 160, 161], is based on the underlying assumption of treatment independence. The underlying assumptions of some existing modelling approaches mean that they will not

be suitable for assessing the treatment sequences for some chronic conditions.

4.5 Taxonomy of Simplifying Assumptions Relating to Treatment-Sequencing Effects

The taxonomy of simplifying assumptions (Table 3) provides a unique and important resource to inform future practice and has the potential to be an important tool for clarifying the extent to which treatment-sequencing effects have been accounted for within a decision model. It can be used as a checklist by modellers to help them consider whether treatment sequencing should be modelled, and what implicit assumptions they may be making. It can also be used by reviewers or policy decision makers to appraise or better understand an existing model. However, to apply the taxonomy, better reporting of the simplifying assumptions made is required.

Our taxonomy focused on the simplifying assumptions made regarding the initial treatment effect (of discrete treatments conditional on their position in the treatment

Table 2 Summary of the different modelling approaches used and their advantages and disadvantages for evaluating treatment sequences

Description of modelling approach	How treatment sequences are conceptualised in the model	Further attributes of the decision problem captured in the model	Advantages	Disadvantages	Included models*
Cohort-based models: *Simulates a closed group of individuals*					
Deterministic decision tree (DT)					
Depicts all possible pathways (series of decisions with associated probabilities and outcomes) over a set period	Treatment sequences were implemented as: (1) A single DT (e.g. initial node representing choice between 1st-line treatments, and subsequent chance nodes response to treatment leading to either a terminal branch (continued treatment) or 2nd-line treatment) (2) Separate DT for each sequence (treatment algorithm) DT structure (and specified timing of chance nodes) also used to account for the following: differential use of subsequent treatment depending on reason for discontinuation; successful treatment (and its discontinuation) leads to either permanent resolution or recurrence requiring re-treatment; multiple level of treatment response (with different time to progression/mortality risk); occurrence of toxic death and all-cause mortality with differential probability and timing; fixed period of treatment administration; some treatments can be skipped	Relapse treated with a previous successful treatment. Duration of response differs according to levels of response. Reason for discontinuation impacts selection of subsequent treatments. Some treatments administered for a fixed period only whilst others continued until progression. Toxic death and all-cause mortality have different probabilities and timing	Can be relatively straightforward to develop and not computationally intensive. Can be easy to interpret and transparent. Can be used as an adjunct or in conjunction with other methods (e.g. Markov cohort, NICE CG152; or partitioned survival, NICE CG131 below)	No explicit time component; governed by fixed timing of outcomes and events (e.g. recurrence or intolerance to treatment). Has finite time horizon; can become exponentially complex (*or bushy*) with additional events and disease states, or extensive treatment sequences. Cannot handle looping/recurring events (reflecting chronic diseases that evolve over time) easily; becomes cumbersome and inefficient when time horizon is long. Poorly suited for complex scenarios/sequences	Dranitsaris model NICE CG81 NICE CG152 Sciatica model Frankum model Knoester model
Stochastic decision tree**					
A type of decision tree that allows for parameter uncertainty	Same as decision tree	Same as decision tree. Not all patients receive all treatments in sequence (*implemented in conjunction with partitioned survival, NICE CG131*).	Same properties as decision tree	Same properties as decision tree	Advanced Simulation Model Greenhalgh model NICE CG131

Table 2 (continued)

Description of modelling approach	How treatment sequences are conceptualised in the model	Further attributes of the decision problem captured in the model	Advantages	Disadvantages	Included models*
Markov cohort					
Simulation of a hypothetical cohort through a set of heath states over time, which is divided into equal intervals (cycles). Involves time-dependent transition between states. In Markov chain transition probabilities are constant over time	Treatment sequences implemented using three different approaches: (1) A series of treatment-specific Markov states (2) As a series of treatment-specific states (or treatment lines) along with additional temporary states representing e.g. adverse effects, relapse (3) As a Markov cycle tree***, with Markov states used to represent different levels of disease activity or natural history. Markov cohort was used for decision problems that related to: the comparison of predefined sequences; or assessing all conceivable sequences (*ordering treatments according to net-benefit/cost-effectiveness per unit time*) (Fig. 2)	Not all patients receive all treatments in the sequence; reflecting reality where patients may skip some treatments. Duration of response differs according to levels of response and treatment line. Probability of continuing treatment or developing toxicity varies with time and for individual treatments (*using tunnel states; or making treatment duration dependent on response level, and the response treatment and line dependent, assuming homogeneous response across treatments but not line of therapy*). Reason for treatment discontinuation impacts selection of subsequent treatments. Some treatments administered for fixed period only; patients in remission may withdraw from treatment. Cross-resistance from previous treatment (*using tunnel states to account for previous treatment*). Monitoring different patient subgroups (*using tunnel states*). *Cycle trees used to account for:* Consequence of adverse effects Different levels of treatment response Some patients continue treatment despite not achieving full/clinical response Changes in disease activity Natural history of the disease Complex treatment pathways	Can be relatively straightforward to construct and communicate; but less so with more complex model structures. Has a time component; events (e.g. relapse or treatment switching) can occur at any time. Allows looping/recurring events. Transitions can be unidirectional or bidirectional. Can be used in conjunction with decision tree (cycle tree) to allow for different treatment response, fluctuating disease activity, complex treatment pathways, etc. The use of cloned subtrees enables ease of update	Markov assumption (*memoryless*): prohibits transition probabilities being dependent on patient's history: time spent in the state, or previous states visited. (*Can be overcome using additional temporary or tunnel states, e.g. making transition dependent on disease progression, and stratification, e.g. response to each treatment line with and without adverse effects.*) Patients can only be in one state at a time; implementing sequences as a series of treatment-specific states only (1) does not allow for additional factors (e.g. different reasons for treatment discontinuation). Cannot account for multiple events within one cycle (e.g. toxicity and progression; *can be overcome using short cycles or cycle trees*). Transitions limited to fixed intervals defined by cycle length. Occurrence of events assumed to be constant over time (Markov chain). Exponential complexity with increasing number of states; modelling extensive treatment sequences (multiple lines) can lead to state explosion, especially when also accounting for additional factors	Albert model Maetzel model Welsing model Tanno model York psoriasis model Beard model Cameron model Davies model Heeg cancer model Lee model Orme model Sawyer model NICE CG137 Shepherd model Smith model Soini model Tebas model Wong model Wu model

Table 2 (continued)

Description of modelling approach	How treatment sequences are conceptualised in the model	Further attributes of the decision problem captured in the model	Advantages	Disadvantages	Included models*
Semi-Markov cohort					
Incorporates the use of a multiple-dimension transition matrix. Assumes transition probabilities depend on the current state, and the time spent in each state depends upon the current and next state	Treatment sequences implemented as a series of treatment-specific health states	Probability of treatment failure decreases with time spent on a specific drug	*Same properties as Markov cohort.* Reduced impact of Markovian assumption (not memoryless; incorporates time dependency). Multidimensional matrix can potentially allow model to reflect patient history or previous treatments (but no study included this)	Patients can only be in one state at a time. Transitions can only occur at fixed intervals defined by cycle length. Only one transition allowed per cycle. Becomes more complex with added states	York epilepsy model
Partitioned survival					
Simulation of a hypothetical cohort through a set of exhaustive and mutually exclusive heath states over time. Time spent in each health state calculated from the area under the curve of survival functions	Treatment sequences implemented as a series of treatment-specific health states	Decreasing probability of remaining on a given treatment with time	Can be relatively straightforward to develop and not computationally intensive. Non-data intensive. Area under the curve can be calculated continuously over time; no cycles required. Can be used in conjunction with decision tree. Can account for extensive treatment sequences (multiple treatment lines)	Cannot account for complex treatment-sequencing algorithms or additional attributes (e.g. adverse effects, disease duration). Underlying structural assumption—surrogate outcomes, e.g. progression free survival, are independent of overall survival	Schadlich model, NICE CG131 (*with decision tree*), Hind model
Individual sampling models: *Simulates one individual at a time. Tracks the past health states of each individual and stochastically models the risk of future events*					
State transition model					
Simulates each individual through a set of exhaustive and mutually exclusive heath states over time, which is advanced in fixed intervals	Fixed treatment sequences modelled; disease activity monitored for each individual over time. Health states generally represented response or non-response to each successive treatment, with the addition of adverse effects as a separate state in some models. Most models also allocated patients to different levels of response after initiating treatment, with those achieving a specific threshold (remission) continuing treatment	Duration of response differs according to levels of response. Changes in disease activity (assumed to relate to level of response, not treatment). Patients follow different disease courses, which cannot be predicted at the onset. Stepped care approach (treatment algorithms)	Not limited by Markov assumption (eliminating need for excessive number of states). A large number of characteristics can be ascribed to individually simulated patients. Access to individual patient data enabled key parameters and events in patient histories to be calculated using multivariate regression, allowing adjusting for important covariates. Can account for heterogeneous population	Transition limited to fixed intervals defined by cycle length. Cannot account for multiple events in one cycle. Can be computationally intensive	Sheffield etanercept model, Bansback model, Sheffield BSRBR model, Diamantpoulus model, Sheffield AHRQ model, Kielhorn model, Kobelt model, Holmes model

Table 2 (continued)

Description of modelling approach	How treatment sequences are conceptualised in the model	Further attributes of the decision problem captured in the model	Advantages	Disadvantages	Included models*
Discrete event situation (DES)					
Simulates time to an event and subsequent events for each individual. Probability of the occurrence and timing of an event is determined by random sampling of a probability distribution. Simultaneously varies multiple variables; inputs vary over time	Treatment sequences implemented in three different ways: (1) Fixed treatment sequences (2) Random selection of pre-defined sequences (3) Developed as part of the modelling process by selecting individual drugs, using a random process, at specific points in the sequence	*When conceived as a simple structure:* Variable time to quitting treatment. Duration of response differs according to levels of response Changes in disease activity (over time and on treatment) Reason for discontinuation impacts selection of subsequent treatments *When developed as a more complex structure:* Treatment selection and cessation based on algorithms reflecting specific clinical guidelines (accounting for treatment eligibility) Unpredictable nature of disease progression Multiple treatment outcomes Different levels of response with partial/complete response leading to treatment withdrawal in some patients Not all patients go on to receive subsequent treatments in the sequence Differential treatment selection for subgroups	Can ascribe a large number of characteristics to individually simulated patients. Can account for heterogeneous population. Not limited by the use of fixed time advancement (cycles). Treatment duration can be modelled as a continuous distribution, specific to individual treatments. Patients can simultaneously be in multiple states, and experience different events. Allows for modelling of complex scenarios and treatment algorithms. Can be computationally more efficient than state transition model. Can be easily adapted to incorporate additional events or patient attributes	Extensive data required including time to event, which may be limited, e.g. time to treatment withdrawal due to adverse effects. Individual patient-level data preferred, but can be based on aggregate data. Model structures can be difficult to communicate and interpret. Computationally challenging in terms of model design and running it	BPM/BRAM Tran-Duy model Lindgren model Birmingham epilepsy model Denis model Heeg schizophrenia model

Table 2 (continued)

Description of modelling approach	How treatment sequences are conceptualised in the model	Further attributes of the decision problem captured in the model	Advantages	Disadvantages	Included models*
Discretely integrated condition event (DICE) simulation					
Decision problem conceptualised as a set of conditions (aspects that persist over time) and events (aspects that occur at a point in time) within spreadsheet tables that specify condition values and event consequences. Provides a single template for implementing a variety of model types (DES, Markov models or hybrids; cohort or individual sampling models; stochastic or deterministic)	Fixed treatment sequences modelled. Events included lack of response and loss of response that lead to treatment initiation. Whilst on treatment patients could experience adverse effects (monitored)	Patient characteristics (baseline risk and response (status, duration) impacts disease milestones. Variable time to quitting treatment. Changes to disease activity (assumed to relate to level of response). Treatment switching based on clinical rules (disease severity and response). Probability of switching treatments in later lines reduced	Can ascribe a large number of characteristics to individually simulated patients. Can account for heterogeneous population. Flexible, allowing combinations of state-transition and time-to-event components in a single model. Treatment duration can be modelled as a continuous distribution, specific to individual treatments. Allows for modelling of complex scenarios and treatment algorithms. Transparent (specifications tabulated rather than programmed in code). Easily modified	Extensive data required including time to event, which may be limited, e.g. time to treatment withdrawal due to adverse effects. Executing simulation using macro in spreadsheet can be slow for complex models	HAS RA model Deniz model
(Open) population-based models: *Allows new cohorts to enter over time*					
Non-terminating population based simulation					
Individual sampling model (DES)	Pre-specified clinical thresholds used to invoke escalation to next treatment	Dynamic disease process (dynamic equations used to project haemoglobin levels over time)	*Has the same properties as DES.* No clear advantage of using open model over other approaches identified for evaluating treatment sequences. However, Cardiff model provides an example of a model that is continually being developed and updated and capable of running using various levels or types of data	*Has the same properties as DES.*	Cardiff T2DM model
Markov multi-cohort model					
Cohort model (Markov)	Markov cycle tree (Markov states represented individual treatments and 'switching' [entry/exit state])	Impact of adding a new drug on health care budget assessed using prevalence approach (target population kept constant over time—entry of newly diagnosed cohort at each cycle)	*Has the same properties as Markov cohort.* No clear advantage of using open model over other approaches identified for evaluating treatment sequences	*Has the same properties as Markov cohort*	Launois model

BPM Birmingham preliminary model, *BRAM* Birmingham Rheumatoid Arthritis Model, *CG* clinical guidelines, *NICE* National Institute for Health and Care Excellence, *T2DM* type 2 diabetes mellitus

*A list of included modelling studies is provided in Online Resource 3 (see the electronic supplementary material)

**Base-case model analysed using Monte Carlo simulation

***Markovian cycle tree (or Markov decision tree) is where events that can occur within a cycle are modelled as a series of chance nodes[183, 184]

Table 3 Taxonomy of simplifying assumptions relating to treatment-sequencing effects used by studies included in the review

Simplifying assumptions taxonomy	
Treatment independence	Treatment effect is *independent of positioning* in treatment sequence
	Treatment effect is dependent on the *number of previous treatments* used, but independent of the type of treatments used
Substitution with another treatment effect	Treatment effect is the same as an alternative treatment from the same class, or a generic class effect—irrespective of positioning in the sequence (*generic effect*)
	Treatment effect is the same as an alternative treatment from the same class, or a generic class effect—matching the same position in the sequence (*positional generic effect*)
	Treatment effect is the same as an alternative (substitute) treatment from a different class of treatments, used at the same point in the sequence (*substitute treatment*)
Modification of treatment effect	Treatment effect is reduced/increased, in line with a multiplier (*multiplication factor*), when used at a later point in the sequence. (Here, the specific multiplication or reduction factor used to *modify* the effect is informed by the available evidence that is also relevant to the treatment of interest.)
	Treatment effect decrements by the same pre-set amount with each successive treatment (*decrementing effect*). (Here, the same generic proportional reduction, used to represent the diminishing effects, is applied at each point in the sequence irrespective of the treatment used. The proportion is not necessarily based on a specific evidence base.)
	Treatment effect is *reduced with disease duration*, and treatments are not as effective when they are used in late disease
Impact of time since previous treatment	Treatment effect is not affected by previous treatments if patients have been in *long-term remission,* and thus can re-use the same treatment(s)/class of treatment(s) as that which achieved the prior remission
Displacement effect ignored	A single treatment effect does not differ when it is *displaced* (i.e. its position in the sequence is changed) by the addition of a new prior treatment (*displacement ignored*)
The use of uncontrolled/observational studies without bias adjustment	*Uncontrolled* trials or *observational studies* provide an un-biased estimate of treatment (sequencing) effects
	Expert consensus provides an un-biased estimate of treatment-sequencing effects

pathway). This incorporates the impact of previous treatment, differential reason for discontinuing previous treatment, and increasing disease duration. However, the taxonomy did not consider the assumptions made about the long-term effect of treatment. Many treatments of chronic conditions, such as rheumatoid arthritis, result in an initial, short-term improvement, followed by a period of waning effect. In some models, when patients move quickly through the sequence of treatments (for example, early discontinuation due to adverse effects), simulated patients can actually benefit from having multiple 'short-term' benefits from different treatments, thus gaining an additive effect. Some included models of inflammatory arthritis attempted to overcome this problem by introducing a 'rebound' effect, which automatically returns the patient to their starting severity (used in, for example, the Diamantopoulos model [89]), or following some natural, background increase (as used in the

BRAM [158]). Although the evidence to support this type of assumption is weak, it is arguably better than the false benefits generated by models otherwise. Similarly, the issue of accumulating short-term benefit can also be problematic where there is an asymmetry in the sequences being compared, for example, the 'adding' decision problem illustrated in Box 2. A false benefit can be introduced when modelling a sequence plus new treatment, in comparison to the model without the added treatment, simply by allowing more 'short-term' effects of treatments.

4.6 Recommendations for Research

An important outcome of the review is the gaps identified in the research evidence. More research is needed to establish when it is necessary to evaluate treatment sequences, and how best to make this decision. This is likely to be a

Table 4 Summary of the frequency of use of the simplifying assumptions

Simplifying assumption used	Total (n = 68)*	Rheumatology studies (n = 36)	Non-rheumatology studies (n = 32)
Treatment independence			
Independent of positioning (IP)	49 (72%)	25 (71%)	24 (75%)
Dependent on number of previous treatments used (NPT)	29 (44%)	19 (54%)	10 (32%)
NPT used in conjunction with IP		*11/19*	*6/8*
Substitution with another treatment effect			
Generic effect (GE)	16 (24%)	14 (40%)	2 (6%)
Positional generic effect (PGE)	17 (26%)	15 (43%)	2 (6%)
Substitute treatment (ST)	1 (2%)	1 (3%)	–
Modification of treatment effect			
Multiplication factor (MF)	10 (15%)	7 (19%)	3 (10%)
Decrementing effect (TD)	4 (6%)	1 (3%)	3 (10%)
Reduced with disease duration (RDD)	4 (6%)	3 (9%)	1 (3%)
Impact of time since previous treatment			
Long-term remission (LR)	2 (3%)	–	2 (6%)
Displacement effect ignored			
Displacement ignored (DI)	28 (42%)	23 (66%)	5 (16%)
The use of uncontrolled/observational studies without bias adjustment (internal validity)**			
Uncontrolled trials or observational studies (UOBS)	20 (30%)	16 (35%)	4 (13%)
Expert consensus (EXC)	4 (6%)	1 (3%)	3 (10%)

CG clinical guidelines, *NICE* National Institute for Health and Care Excellence

*Four modelling studies included in the review of modelling approaches were not included in the review of simplifying assumptions. These include two (NICE CG131; Hind, 2008: NICE TA79) [53, 56] that obtained data on clinical effectiveness from sequencing trials and two (McEwan et al., 2010, and Launois et al., 2008) [102, 108] that did not evaluate the clinical effectiveness of treatment sequences, but were included as they provided relevant examples of specific modelling techniques

**This relates to no adjustment made to issues relating to the use of observational data rather than data form randomised controlled trials

condition-specific endeavour, but the methods will be relevant across different clinical scenarios.

Further research is needed to identify how best to develop a summary treatment effect of whole sequences or discrete interventions conditional on positioning in the sequence. This requires improved reporting on previous and subsequent treatment within primary studies, including better data on reasons for discontinuing or switching treatment. Access to individual patient-level data is also key here [35, 162].

Real-world disease-specific data sources can provide essential follow-up data on entire treatment sequences, and potentially be used to emulate a pragmatic randomised trial of dynamic treatment sequences [27, 163–165]. If these data sources are going to be useful, treatment sequences need to be considered during the planning and development stages. They will also need to go through many high-quality validation studies [164]. The evaluation of whole treatment sequences using real-world data also needs to take into account the potential biases listed in Box 1.

Finally, little reference was made within existing research on the potential, or actual role, of incorporating patient perspectives into the evaluation of treatment sequences. Further work is needed to develop the optimal approach for involving members of the public in HTA of treatment sequences, which should be informed by existing guidance and recent research on patient and public involvement in systematic reviews and economic evaluations [166–171]. As experienced-based experts, patients can contribute essential knowledge that is complementary to that of other key stakeholders, such as clinicians and policy makers. Their involvement, on an equal basis to other stakeholders, is likely to be relevant to all stages of the HTA, including refining the scope and decision problem, the evidence synthesis, evidence interpretation and integration, and dissemination and application [172].

5 Conclusions

The review illuminates a significant gap in methods development. It also demonstrates important limitations in the primary studies, which tended to focus on the evaluation of discrete treatments, with poor reporting of any previous or subsequent treatments. The increasing use of NMA in HTA demonstrates an acknowledgment that clinical and policy decision making should account for the multiple treatments available for many chronic conditions. However, the sequential use of these treatments has yet to be accounted for within clinical evaluations, with most meta-analysis being conducted of discrete treatments that may or may not be stratified by line of therapy. The economic modelling exposes the need to consider treatment sequences, but this is often based on the simplifying assumption of treatment

independence. This can lead to misrepresentation of the true level of uncertainty, potential bias in estimating the effectiveness and cost-effectiveness of treatments, and eventually the wrong decision.

In summary, there has been no co-ordinated approach to the important issue of evaluating the effectiveness and cost-effectiveness of treatment sequencing. This is a major shortfall at a time when the cohort of people with complex chronic conditions, requiring sequential treatments, is increasing. The findings of the review will help policy makers and researchers gain traction in answering questions about the effectiveness of different treatment sequences.

Declarations

Conflict of interest RL, DH, AJS, and CW have no conflicts of interest that are directly relevant to the content of this article.

Ethical approval This article does not contain any studies with animals or humans performed by any of the authors.

Consent to participate Not applicable.

Consent for publication Not applicable.

Code availability Not applicable.

Author contributions All authors contributed to the conception and design of the review. RAL performed the literature search and data analysis and drafted the original manuscript. DH and AJS provided input regarding interpreting the results and critically revised the work. All authors read and approved the final manuscript. RAL will act as the overall guarantor.

References

1. Ioannidis JPA, Patsopoulos NA, Rothstein HR. Reasons or excuses for avoiding meta-analysis in forest plots. BMJ. 2008;336:1413. https://doi.org/10.1136/bmj.a117.

2. Viola MG, Diamantopoulos A. Determining the baseline strategy in a cost-effectiveness analysis with treatment sequences. Appl Health Econ Health Policy. 2020;18(1):17–29. https://doi.org/10.1007/s40258-019-00514-2.

3. Zheng Y, Pan F, Sorensen S. Modeling treatment sequences in pharmacoeconomic models. Pharmacoeconomics. 2017;35(1):15–24. https://doi.org/10.1007/s40273-016-0455-3.

4. Tosh J, Stevenson M, Akehurst R. Health economic modelling of treatment sequences for rheumatoid arthritis: a systematic review. Curr Rheumatol Rep. 2014;16(10):447. https://doi.org/10.1007/s11926-014-0447-2.

5. Messner DA. Evaluating the comparative effectiveness of treatment sequences in oncology: a new approach. J Comp Eff Res. 2015;4(6):537–54. https://doi.org/10.2217/cer.15.47.

6. Kamerow D. How can we treat multiple chronic conditions? BMJ. 2012;344:e1487. https://doi.org/10.1136/bmj.e1487.

7. Chakraborty B. Dynamic treatment regimes for managing chronic health conditions: a statistical perspective. Am J Public Health. 2011;101(1):40–5. https://doi.org/10.2105/AJPH.2010.198937.

8. Mauskopf J. A methodological review of models used to estimate the cost effectiveness of antiretroviral regimens for the treatment of HIV infection. Pharmacoeconomics. 2013;31(11):1031–50.

9. Lloyd S, Bujkiewicz S, Wailoo AJ, Sutton AJ, Scott DA. The effectiveness of anti-TNF-a therapies when used sequentially in rheumatoid arthritis patients: a systematic review and meta-analysis. Rheumatology. 2010;49:2313–21.

10. Christensen AW, Tarp S, Furst DE, Døssing A, Amris K, Bliddal H, et al. Most trial eligibility criteria and patient baseline characteristics do not modify treatment effect in trials using targeted therapies for rheumatoid arthritis: a meta-epidemiological study. PLoS ONE. 2015;10(9):e0136982. https://doi.org/10.1371/journal.pone.0136982.

11. Stenner F, Chastonay R, Liewen H, Haile SR, Cathomas R, Rothermundt C, et al. A pooled analysis of sequential therapies with sorafenib and sunitinib in metastatic renal cell carcinoma. Oncology. 2012;6:333–40.

12. Jamnitski A, Bartelds GM, Nurmohamed MT, van Schouwenburg PA, van Schaardenburg D, Stapel SO, et al. The presence or absence of antibodies to infliximab or adalimumab determines the outcome of switching to etanercept. Ann Rheum Dis. 2011;70(2):284–8. https://doi.org/10.1136/ard.2010.135111.

13. Oronsky B, Carter CA, Reid TR, Scicinski J, Oronsky A, Lybeck M, et al. Confirmatory trials in the evaluation of anticancer medicinal products in man–PFS2: a measure of therapeutic action-at-a-distance. Neoplasia. 2015;17(9):716–22. https://doi.org/10.1016/j.neo.2015.09.001.

14. Albert DA, Aksentijevich S, Hurst S, Fries JF, Wolfe F. Modeling therapeutic strategies in rheumatoid arthritis: use of decision analysis and Markov models. J Rheumatol. 2000;27(3):644–52.

15. National Institute for Health and Care Excellence. Breast cancer (advanced): NICE clinical guideline 81. London: NICE NICE technology appraisal guidance; 2009.

16. Rodgers M, Epstein D, Bojke L, Yang H, Craig D, Fonseca T, et al. Etanercept, infliximab and adalimumab for the treatment of psoriatic arthritis: a systematic review and economic evaluation. Health Technol Assess. 2011;15(10):1–329. https://doi.org/10.3310/hta15100.

17. Wilby J, Kainth A, Hawkins N, Epstein D, McIntosh H, McDaid C, et al. Clinical effectiveness, tolerability and cost-effectiveness of newer drugs for epilepsy in adults: a systematic review and economic evaluation. Health Technol Assess. 2005;9(15):1–157.

18. National Institute for Health and Care Excellence. Epilepsy: NICE clinical guideline 137. London: NICE; 2012a.

19. Kalden JR, Schulze-Koops H. Immunogenicity and loss of response to TNF inhibitors: implications for rheumatoid arthritis treatment. Nat Rev Rheumatol. 2017;13(12):707–18.

20. Dimopoulos G, Siempos II, Korbila IP, Manta KG, Falagas M. Comparison of first-line with second-line antibiotics for acute exacerbations of chronic bronchitis: a metaanalysis of randomized controlled trials. Chest. 2007;132(2):447–55.

21. Eichelberg C, Vervenne WL, De Santis M, Fischer von Weikersthal L, Goebell PJ, Lerchenmüller C et al. SWITCH: a randomised, sequential, open-label study to evaluate the efficacy and safety of sorafenib-sunitinib versus sunitinib-sorafenib in the treatment of metastatic renal cell cancer. Eur Urol. 2015;68(5):837–47. 10.1016/j.eururo.2015.04.017

22. Hider SL, Buckley C, Silman AJ, Symmons DP, Bruce IN. Factors influencing response to disease modifying antirheumatic drugs in patients with rheumatoid arthritis. J Rheumatol. 2005;32(1):11–6.

23. Anderson JJ, Wells G, Verhoeven AC, Felson DT. Factors predicting response to treatment in rheumatoid arthritis: the importance of disease duration. Arthritis Rheumatol. 2000;43(1):22–9.

24. Allen F, Montgomery S, Maruszczak M, Kusel J, Adlard N. Convergence yet continued complexity: a systematic review and critique of health economic models of relapsing-remitting multiple sclerosis in the United Kingdom. Value Health. 2015;18(6):925–38. https://doi.org/10.1016/j.jval.2015.05.006.

25. Hawkins N, Sculpher M, Epstein D. Cost-effectiveness analysis of treatments for chronic disease: using R to incorporate time dependency of treatment response. Med Decis Mak. 2005;25(5):511–9.

26. Almirall D, Nahum-Shani I, Sherwood NE, Murphy SA. Introduction to SMART designs for the development of adaptive interventions: with application to weight loss research. Transl Behav Med. 2014;4(3):260–74. https://doi.org/10.1007/s13142-014-0265-0.

27. Lei H, Nahum-Shani I, Lynch K, Oslin D, Murphy SA. A "SMART" design for building individualized treatment sequences. Annu Rev Clin Psychol. 2012;8:21–48. https://doi.org/10.1146/annurev-clinpsy-032511-143152.

28. Kidwell KM, Postow MA, Panageas KS. Sequential, multiple assignment, randomized trial designs in immuno-oncology research. Clin Cancer Res. 2018;24(4):730–6. https://doi.org/10.1158/1078-0432.CCR-17-1355.

29. Almirall S, Collins LM, Murphy S (2011) Introduction to adaptive interventions and SMART designs. In: Getting SMART about developing individualized sequences of health interventions. CPDD, Hollywood, Florida. shttps://people.seas.harvard.edu/~samurphy/seminars/NCDEUAlmirall.pdf. Accessed 4th Nov 2013

30. Kidwell KM. SMART designs in cancer research: past present and future. Clin Trials. 2014;11(4):445–56. https://doi.org/10.1177/1740774514525691.

31. Murphy SA, Lynch KG, Oslin D, McKay JR, TenHave T. Developing adaptive treatment strategies in substance abuse research. Drug Alcohol Depend. 2007;88(Suppl 2):S24–30. https://doi.org/10.1016/j.drugalcdep.2006.09.008.

32. Sutton AJ, Higgins JPT. Recent developments in meta-analysis. Stat Med. 2008;27:625–50.

33. Efthimiou O, Debray TP, van Valkenhoef G, Trelle S, Panayidou K, Moons KG, et al. GetReal in network meta-analysis: a review of the methodology. Res Synth Methods. 2016;7(3):236–63. https://doi.org/10.1002/jrsm.1195.

34. Efthimiou O, Mavridis D, Debray TP, Samara M, Belger M, Siontis GC, et al. Combining randomized and non-randomized evidence in network meta-analysis. Stat Med. 2017;36(8):1210–26. https://doi.org/10.1002/sim.7223.

35. Debray TP, Moons KG, van Valkenhoef G, Efthimiou O, Hummel N, Groenwold RH, et al. Get real in individual participant data (IPD) meta-analysis: a review of the methodology. Res Synth Methods. 2015;6(4):293–309. https://doi.org/10.1002/jrsm.1160.

36. Guise J-M, Chang C, Viswanathan M, Glick S, Treadwell JR, Umscheid CA et al. Systematic Reviews of Complex Multicomponent Health Care Interventions. Rockville (MD): Agency for Healthcare Research and Quality (US); 2014.

37. Panayidou K, Gsteiger S, Egger M, Kilcher G, Carreras M, Efthimiou O, et al. GetReal in mathematical modelling: a review of studies predicting drug effectiveness in the real world. Res Synth Methods. 2016;7(3):264–77. https://doi.org/10.1002/jrsm.1202.

38. Tosh J, Trowman R. Briefing paper for methods review working party on treatment sequences and downstream costs. Briefing paper for the update of the NICE methods guide 2011.

39. Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovitch C, Song F et al. Evaluating non-randomised intervention studies. Health Technol Assess. 2003;7(27).

40. Kaltenthaler E, Tappenden P, Paisley S, Squires H. NICE DSU Technical Support Document 13: Identifying and Reviewing Evidence to Inform the Conceptualisation and Population of Cost-Effectiveness Models. NICE Decision Support Unit Technical Support Documents. Sheffield: NICE Decision Support Unit, ScHARR; 2011.

41. Gentles SJ, Charles C, Nicholas DB, Ploeg J, McKibbon KA. Reviewing the research methods literature: principles and strategies illustrated by a systematic overview of sampling in qualitative research. Syst Rev. 2016;5(1):172.

42. Dixon-Woods M, Bonas S, Booth A, Jones DR, Miller T, Shaw RL, et al. How can systematic reviews incorporate qualitative research? A critical perspective. Qual Res. 2006;6(1):27–44. https://doi.org/10.1177/1468794106058867.

43. Booth A. Searching for qualitative research for inclusion in systematic reviews: a structured methodological review. Syst Rev. 2016;5:74. https://doi.org/10.1186/s13643-016-0249-x.

44. Ghabri S, Lam L, Bocquet F, Spath HM. Systematic literature review of economic evaluations of biological treatment sequences for patients with moderate to severe rheumatoid arthritis previously treated with disease-modifying anti-rheumatic drugs. Pharmacoeconomics. 2020;38(5):459–71.

45. Ghabri S, Binard A, Pers YM, Maunoury F, Caro JJ. Economic evaluation of sequences of biological treatments for patients with moderate-to-severe rheumatoid arthritis and inadequate response or intolerance to methotrexate in France. Value Health. 2020;23(4):461–70.

46. Deniz B, Ambavane A, Yang SA, Doan J, Rao S, et al. Treatment sequences for advanced renal cell carcinoma: a health economic assessment. PLoS ONE. 2019;14(8):0215761. https://doi.org/10.1371/journal.pone.0215761.

47. Abrams K, Bujkiewicz S, Dequen P, Jenkins D, Martina R. WP1: Deliverable 1.5 (Case Study Review: Rheumatoid Arthritis).

https://wwwimi-getrealeu/Portals/1/Documents/01%20deliverables/Deliverable%20Report%20D15_Rheumatoid%20Arthritis_websiteversionpdf. 2016; Accessed Dec 2017.

48. Connock M, Frew E, Evans B-W, Bryan S, Cummins C, Fry-Smith A, et al. The clinical effectiveness and cost-effectiveness of newer drugs for children with epilepsy. A systematic review. Health Technol Assess. 2006;10(7):1–118. https://doi.org/10.3310/hta10070.

49. Cooper C, Katona C, Lyketsos K, Blazer D, Brodaty H, Rabins P, et al. A systematic review of treatments for refractory depression in older people. Am J Psychiatry. 2011;168(7):681–8. https://doi.org/10.1176/appi.ajp.2011.10081165.

50. Finnerup NB, Otto M, McQuay HJ, Jensen TS, Sindrup SH. Algorithm for neuropathic pain treatment: an evidence based proposal. Pain. 2005;118(3):289–305.

51. Grothey A, Sargent D, Goldberg RM, Schmoll HJ. Survival of patients with advanced colorectal cancer improves with the availability of fluorouracil-leucovorin, irinotecan, and oxaliplatin in the course of treatment. J Clin Oncol. 2004;22(7):1209–14.

52. Heng DY, Signorovitch J, Swallow E, Li N, Zhong Y, Qin P, et al. Comparative effectiveness of second-line targeted therapies for metastatic renal cell carcinoma: a systematic review and meta-analysis of real-world observational studies. PLoS ONE. 2014;9(12):e114264. https://doi.org/10.1371/journal.pone.0114264.

53. Hind D, Tappenden P, Tumur I, Eggington S, Sutcliffe P, Ryan A. The use of irinotecan, oxaliplatin and raltitrexed for the treatment of advanced colorectal cancer: systematic review and economic evaluation. Health Technol Assess. 2008;12(15):1–162.

54. Kanters S, Druyts E, Mills EJ, Thorlund K. What drives the comparative effectiveness of biologics vs. methotrexate in rheumatoid arthritis? Meta-regression and graphical inspection of suspected clinical factors. Rheumatology 2014;53(7):1264-73. https://doi.org/10.1093/rheumatology/ket492.

55. Mandema JW, Salinger DH, Baumgartner SW, Gibbs MA. A dose-response meta-analysis for quantifying relative efficacy of biologics in rheumatoid arthritis. Clin Pharmacol Ther. 2011;90:828–35. https://doi.org/10.1038/clpt.2011.256.

56. National Institute for Health and Care Excellence. Colorectal cancer: NICE clinical guideline 131. London: NICE; 2011.

57. Nixon RA, Bansback N, Brennan A. Using mixed treatment comparisons and meta-regression to perform indirect comparisons to estimate the efficacy of biologic treatments in rheumatoid arthritis. Stat Med. 2007;26:1237–54.

58. Rendas-Baum R, Wallenstein GV, Koncz T, Kosinski M, Yang M, Bradley J et al. Evaluating the efficacy of sequential biologic therapies for rheumatoid arthritis patients with an inadequate response to tumor necrosis factor-a inhibitors. https://arthritis-researchcom/content/13/1/R25. 2011;13:R25. doi:https://doi.org/10.1186/ar3249.

59. Ruhé HG, Huyser J, Swinkels JA, Schene AH. Switching antidepressants after a first selective serotonin reuptake inhibitor in major depressive disorder: a systematic review. J Clin Psychiatry. 2006;67(12):1836–55.

60. Salliot C, Finckh A, Katchamart W, Lu Y, Sun Y, Bombardier C, et al. Indirect comparisons of the efficacy of biological antirheumatic agents in rheumatoid arthritis in patients with an inadequate response to conventional disease-modifying antirheumatic drugs or to an anti-tumour necrosis factor agent: a meta-analysis. Ann Rheum Dis. 2011;70(2):266–71. https://doi.org/10.1136/ard.2010.132134.

61. Schmitz S, Adams R, Walsh CD, Barry M, FitzGerald O. A mixed treatment comparison of the efficacy of anti-TNF agents in rheumatoid arthritis for methotrexate non-responders demonstrates differences between treatments: a Bayesian approach. Ann

Rheum Dis. 2012;71(2):225–30. https://doi.org/10.1136/annrheumdis-2011-200228.

62. Schoels M, Aletaha D, Smolen JS, Wong JB. Comparative effectiveness and safety of biological treatment options after tumour necrosis factor α inhibitor failure in rheumatoid arthritis: systematic review and indirect pairwise meta-analysis. Ann Rheum Dis. 2012;71:1303–8.

63. Singh JA, Christensen R, Wells GA, Suarez-Almazor ME, Buchbinder R, Lopez-Olivo MA, et al. A network meta-analysis of randomized controlled trials of biologics for rheumatoid arthritis: a Cochrane overview. Can Med Assoc J. 2009;181(11):787–96. https://doi.org/10.1503/cmaj.091391.

64. Suarez-Almazor M, Ortiz Z, Lopez-Olivo M, Pak C, Skidmore B, Kimmel B et al. Infliximab and etanercept in rheumatoid arthritis: timing, dose escalation, and switching [Technology report no 86] Ottawa: Canadian Agency for Drugs and Technologies in Health. 2007.

65. Ades AE, Madan J, Welton NJ. Indirect and mixed treatment comparisons in arthritis research. Rheumatology (Oxford). 2011;50 (Suppl 4):iv5-9. https://doi.org/10.1093/rheumatology/ker241.

66. Lewis RA, Williams NH, Sutton AJ, Burton K, Din NU, Matar HE, et al. Comparative clinical effectiveness of management strategies for sciatica: systematic review and network meta-analyses. Spine J. 2015;15(6):1461–77. https://doi.org/10.1016/j.spinee.2013.08.049.

67. Clark W, Jobanputra P, Barton P, Burls A. The clinical and cost-effectiveness of anakinra for the treatment of rheumatoid arthritis in adults: a systematic review and economic analysis. Health Technol Assess. 2004;8(18):1–105.

68. Tran-Duy A, Boonen A, van de Laar MAFJ, Franke AC, Severens JL. A discrete event modelling framework for simulation of long-term outcomes of sequential treatment strategies for ankylosing spondylitis. Ann Rheum Dis. 2011;70(12):2111–8. https://doi.org/10.1136/annrheumdis-2011-200333.

69. Schadlich PK, Zeidler H, Zink A, Gromnica-Ihle E, Schneider M, Straub C et al. Modelling cost effectiveness and cost utility of sequential DMARD therapy including leflunomide for rheumatoid arthritis in Germany: II. The contribution of leflunomide to efficiency. Pharmacoeconomics. 2005;23(4):395-420.

70. Holmes CB, Zheng H, Martinson NA, Freedberg KA, Walensky RP. Optimizing treatment for HIV-infected South African women exposed to single-dose nevirapine: balancing efficacy and cost. Clin Infect Dis. 2006;42(12):1772–80.

71. Russell A, Beresniak A, Bessette L, Haraoui B, Rahman P, Thorne C, et al. Cost-effectiveness modeling of abatacept versus other biologic agents in DMARDS and anti-TNF inadequate responders for the management of moderate to severe rheumatoid arthritis. Clin Rheumatol. 2009;28:403–12.

72. Diamantopoulos A, Benucci M, Capri S, Berger W, Wintfeld N, Giuliani G, et al. Economic evaluation of tocilizumab combination in the treatment of moderate-to-severe rheumatoid arthritis in Italy. J Med Econ. 2012;15(3):576–85. https://doi.org/10.3111/13696998.2012.665110.

73. Gottenberg JE, Brocq O, Perdriger A, Lassoued S, Berthelot J-M, Wendling D et al. Non-TNF-targeted biologic vs a second anti-TNF drug to treat rheumatoid arthritis in patients with insufficient response to a first anti-TNF drug: a randomized clinical trial. JAMA. 2016;316(11):1172–80. https://doi.org/10.1001/jama.2016.13512

74. Anis AH, Bansback N, Sizto S, Gupta SR, Willian MK, Feldman SR. Economic evaluation of biologic therapies for the treatment of moderate to severe psoriasis in the United States. J Dermatolog Treat. 2011;22(2):65–74. https://doi.org/10.3109/09546630903551258.

75. Bansback NJ, Brennan A, Ghatnekar O. Cost effectiveness of adalimumab in the treatment of patients with moderate to severe rheumatoid arthritis in Sweden. Ann Rheum Dis. 2005;64:995–1002.

76. Barton P, Jobanputra P, Wilson J, Bryan S, Burls A. The use of modelling to evaluate new drugs for patients with a chronic condition: the case of antibodies against tumour necrosis factor in rheumatoid arthritis. Health Technol Assess. 2004;8(11):1–91.

77. Beard SM, Roskell N, Le TK, Zhao Y, Coleman A, Ang D, et al. Cost effectiveness of duloxetine in the treatment of fibromyalgia in the United States. J Med Econ. 2011;14(4):463–76. https://doi.org/10.3111/13696998.2011.586389.

78. Bensmail D, Ward AB, Wissel J, Motta F, Saltuari L, Lissens J, et al. Cost-effectiveness modeling of intrathecal baclofen therapy versus other interventions for disabling spasticity. Neurorehabil Neural Repair. 2009;23(6):546–52. https://doi.org/10.1177/1545968308328724.

79. Beresniak A, Ariza-ariza R, Garcia-Llorente JF, Ramirez-Arellano A, Dupont D. Modelling cost-effectiveness of biologic treatments based on disease activity scores for management of rheumatoid arthritis in Spain. International Journal of Inflammation. 2011;v2011:727634. https://doi.org/10.4061/2011/727634.

80. Beresniak A, Baerwald C, Zeidler H, Krüger K, Neubauer AS, Dupont D, et al. Cost-effectiveness simulation model of biologic strategies for treating to target rheumatoid arthritis in Germany. Clin Exp Rheumatol. 2013;31(3):400–8.

81. Brennan A, Bansback N, Reynolds A, Conway P. Modelling the cost-effectiveness of etanercept in adults with rheumatoid arthritis in the UK. Rheumatology. 2004;43(1):62–72. https://doi.org/10.1093/rheumatology/keg451.

82. Brennan A, Bansback N, Nixon R, Madan J, Harrison M, Watson K, et al. Modelling the cost effectiveness of TNF-alpha antagonists in the management of rheumatoid arthritis: results from the British Society for Rheumatology Biologics Registry. Rheumatology. 2007;46(8):1345–54.

83. Cameron DA, Camidge DR, Oyee J, Hirsch M. Economic evaluation of fulvestrant as an extra step in the treatment sequence for ER-positive advanced breast cancer. Br J Cancer. 2008;99(12):1984–90. https://doi.org/10.1038/sj.bjc.6604790.

84. Chen Y-F, Jobanputra P, Barton PM, Bryan S, Clark W, Fry-Smith A, et al. A systematic review of the effectiveness of adalimumab, etanercept and infliximab for the treatment of rheumatoid arthritis in adults and an economic evaluation of their cost-effectiveness. Health Technol Assess. 2006;10(42):1–248.

85. Cimmino MA, Leardini G, Salaffi F, Intorcia M, Bellatreccia A, Dupont D, et al. Assessing the cost-effectiveness of biologic agents for the management of moderate-to-severe rheumatoid arthritis in anti-TNF inadequate responders in Italy: a modelling approach. Clin Exp Rheumatol. 2011;29(4):633–41.

86. Coyle D, Judd M, Blumenauer B, Cranney A, Maetzel A, Tugwell P et al. Infliximab and etanercept in patients with rheumatoid arthritis: a systematic review and economic evaluation. Technology Report Issue 64. Ottawa: Canadian Coordinating Office for Health Technology Assessment (CCOHTA); 2006.

87. Davies A, Vardeva K, Loze J-Y, L'Italien GJ, Sennfalt K, van Baardewijk M. Cost-effectiveness of atypical antipsychotics for the management of schizophrenia in the UK. Curr Med Res Opin. 2008;24(11):3275–85. https://doi.org/10.1185/03007990802507547.

88. Davies A, Cifaldi MA, Segurado OG, Weisman MH. Cost-effectiveness of sequential therapy with tumor necrosis factor antagonists in early rheumatoid arthritis. J Rheumatol. 2009;36(1):16–26. https://doi.org/10.3899/jrheum.080257.

89. Denis P, Le Pen C, Umuhire D, Berdeaux G. Treatment carryover impacts on effectiveness of intraocular pressure lowering agents,

estimated by a discrete event simulation model. Eur J Ophthalmol. 2008;18(1):44–51.

90. Diamantopoulos A, Finckh A, Huizinga T, Sungher DK, Sawyer L, Neto D, et al. Tocilizumab in the treatment of rheumatoid arthritis: a cost-effectiveness analysis in the UK. Pharmacoeconomics. 2014;32(8):775–87. https://doi.org/10.1007/s40273-014-0165-7.

91. Dranitsaris G, Truter I, Lubbe MS, Sriramanakoppa NN, Mendonca VM, Mahagaonkar SB. Using pharmacoeconomic modelling to determine value-based pricing for new pharmaceuticals in Malaysia Malaysian. J Med Sci. 2011;18(4):32–43.

92. Erhardt W, Bergenheim K, Duprat-Lomon I, McEwan P. Cost effectiveness of saxagliptin and metformin versus sulfonylurea and metformin in the treatment of type 2 diabetes mellitus in Germany: a Cardiff diabetes model analysis. Clin Drug Invest. 2012;32(3):189–202. https://doi.org/10.2165/11597060-000000000-00000.

93. Finckh A, Bansback N, Marra CA, Anis AH, Michaud K, Lubin S, et al. Treatment of very early rheumatoid arthritis with symptomatic therapy, disease-modifying antirheumatic drugs, or biologic agents: a cost-effectiveness analysis. Ann Intern Med. 2009;151(9):612–21. https://doi.org/10.7326/0003-4819-151-9-200911030-00006.

94. Fitzsimmons D, Phillips CJ, Bennett H, Jones M, Williams N, Lewis R, et al. Cost effectiveness of different management strategies for sciatica. Pain. 2014;2014(155):1318–27.

95. Frankum LE, Nightengale B, Russo CL, Sarnes M. Pharmacoeconomic analysis of sequential treatment pathways in the treatment of onychomycosis. Manag Care Interface. 2005;18(1):55–63.

96. Greenhalgh J, Knight C, Hind D, Beverley C, Walters S. Clinical and cost-effectiveness of electroconvulsive therapy for depressive illness, schizophrenia, catatonia and mania: systematic reviews and economic modelling studies. Health Technol Assess. 2005;9(9):1–156.

97. Hallinen TA, Soini EJ, Eklund K, Puolakka K. Cost-utility of different treatment strategies after the failure of tumour necrosis factor inhibitor in rheumatoid arthritis in the Finnish setting. Rheumatology. 2010;49(4):767–77.

98. Heeg B, Buskens E, Botteman M, Caleo S, Ingham M, Damen J, et al. The cost-effectiveness of atypicals in the UK. Value in Health. 2008;11(7):1007–21. https://doi.org/10.1111/j.1524-4733.2008.00344.x.

99. Heeg B, van Agthoven M, van Beurden-Tan C, Liwing J, Mellqvist U-H, Plesner T et al. Chapter 5: Treatment Sequencing Survival Model for patients with Multiple Myeloma ineligible for stem cell transplantation (SCT). Developing health economic models of chronic diseases for reimbursement purposes. Groningen: University of Groningen; 2015. 10.1016/j.jval.2014.12.008. https://www.rug.nl/research/portal/files/19289584/Chapter_5.pdf. Accessed Dec 2016

100. Jobanputra P, Barton P, Bryan S, Burls A. The effectiveness of infliximab and etanercept for the treatment of rheumatoid arthritis: a systematic review and economic evaluation. Health Technol Assess. 2002;6(21):1–110.

101. Kielhorn A, Porter D, Diamantopoulos A, Lewis G. UK cost-utility analysis of rituximab in patients with rheumatoid arthritis that failed to respond adequately to a biologic disease-modifying antirheumatic drug. Curr Med Res Opin. 2008;24(9):2639–50. https://doi.org/10.1185/03007990802321683.

102. Knoester PD, Deckers CL, Termeer EH, Boendermaker AJ, Kotsopoulos IA, de Krom MC, et al. A cost-effectiveness decision model for antiepileptic drug treatment in newly diagnosed epilepsy patients. Value Health. 2007;10(3):173–82.

103. Kobelt G, Lekander I, Lang A, Raffeiner B, Botsios C, Geborek P. Cost-effectiveness of etanercept treatment in early active rheumatoid arthritis followed by dose adjustment. Int J Technol Assess Health Care. 2011;27(3):193–200.

104. Launois R, Payet S, Saidenberg-Kermanac'h N, Francesconi C, Riou Franca L, Boissier MC. Budget impact model of rituximab after failure of one or more TNFalpha inhibitor therapies in the treatment of rheumatoid arthritis. Joint Bone Spine. 2008;75(6):688–95.

105. Lee HY, Yang BM, Hong JM, Lee TJ, Kim BG, Kim JW, et al. Cost-utility analysis for platinum-sensitive recurrent ovarian cancer therapy in South Korea: results of the polyethylene glycolated liposomal doxorubicin/carboplatin sequencing model. ClinicoEcon Outcomes Res. 2013;5(1):297–307. https://doi.org/10.2147/CEOR.S42170.

106. Lindgren P, Geborek P, Kobelt G. Modeling the cost-effectiveness of treatment of rheumatoid arthritis with rituximab using registry data from Southern Sweden. Int J Technol Assess Health Care. 2009;25(2):181–9.

107. Lux MP, Hartmann M, Jackisch C, Raab G, Schneeweiss A, Possinger K, et al. Cost-utility analysis for advanced breast cancer therapy in Germany: results of the fulvestrant sequencing model. Breast Cancer Res Treat. 2009;117(2):305–17. https://doi.org/10.1007/s10549-008-0294-9.

108. Maetzel A, Strand V, Tugwell P, Wells G, Bombardier C. Cost Effectiveness of adding leflunomide to a 5-year strategy of conventional disease-modifying antirheumatic drugs in patients with rheumatoid arthritis. Arthritis Rheum. 2002;47(6):655–61.

109. Malottki K, Barton P, Tsourapas A, Uthman AO, Liu Z, Routh K, et al. Adalimumab, etanercept, infliximab, rituximab and abatacept for the treatment of rheumatoid arthritis after the failure of a tumour necrosis factor inhibitor: a systematic review and economic evaluation. Health Technol Assess. 2011;15(14):1–278. https://doi.org/10.3310/hta15140.

110. McEwan P, Evans M, Bergenheim K. A population model evaluating the costs and benefits associated with different oral treatment strategies in people with type 2 diabetes. Diabetes Obes Metab. 2010;12(7):623–30. https://doi.org/10.1111/j.1463-1326.2010.01198.x.

111. Merkesdal S, Kirchhoff T, Wolka D, Ladinek G, Kielhorn A, Rubbert-Roth A. Cost-effectiveness analysis of rituximab treatment in patients in Germany with rheumatoid arthritis after etanercept-failure. Eur J Health Econ. 2010;11(1):95–104. https://doi.org/10.1007/s10198-009-0205-y.

112. National Institute for Health and Care Excellence. Crohn's disease: management: NICE clinical guideline 152. London: NICE; 2012b.

113. Orme M, Collins S, Loftus J. Long-term medical management of primary open-angle glaucoma and ocular hypertension in the UK: optimizing cost-effectiveness and clinic resources by minimizing therapy switches. J Glaucoma. 2012;21(7):433–49. https://doi.org/10.1097/IJG.0b013e31821dac2a.

114. Puolakka K, Blafield H, Kauppi M, Luosujarvi R, Peltomaa R, Leikola-Pelho T, et al. Cost-effectiveness modelling of sequential biologic strategies for the treatment of moderate to severe rheumatoid arthritis in Finland. Open Rheumatol J. 2012;1:38–43.

115. Saraux A, Gossec L, Goupille P, Bregman B, Boccard E, Dupont D, et al. Cost-effectiveness modelling of biological treatment sequences in moderate to severe rheumatoid arthritis in France. Rheumatology. 2010;49(4):733–40. https://doi.org/10.1093/rheumatology/kep434.

116. Sawyer L, Samarasekera EJ, Wonderling D, Smith CH. Topical therapies for the treatment of localized plaque psoriasis in primary care: a cost-effectiveness analysis. Br J Dermatol. 2013;168(5):1095–105. https://doi.org/10.1111/bjd.12261.

117. Schipper LG, Wietske K, den Broeder AA, van der Laar MA, Adang EMM, Fransen J, et al. Treatment strategies aiming at remission in early rheumatoid arthritis patients: starting with

methotrexate monotherapy is cost-effective. Rheumatology. 2011;50:1320–30.

118. Shepherd J, Jones J, Takeda A, Davidson P, Price A. Adefovir dipivoxil and pegylated interferon alfa-2a for the treatment of chronic hepatitis B: a systematic review and economic evaluation. Health Technol Assess. 2006;10(28):iii-iv, xi-xiv, 1-183.

119. Sizto S, Bansback N, Feldman SR, Willian MK, Anis AH. Economic evaluation of systemic therapies for moderate to severe psoriasis. Br J Dermatol. 2009;160(6):1264–72. https://doi.org/10.1111/j.1365-2133.2008.08962.x.

120. Smith KJ, Roberts MS. Sequential medication strategies for postherpetic neuralgia: a cost-effectiveness analysis. J Pain. 2007;8(5):396–404.

121. Soini EJ, Martikainen JA, Vihervaara V, Mustonen K, Nousi-ainen T. Economic evaluation of sequential treatments for follicular non-hodgkin lymphoma. Clin Ther. 2012;34(4):915-25. e2. https://doi.org/10.1016/j.clinthera.2012.02.019.

122. Tanno M, Nakamura I, Ito K, Tanaka H, Ohta H, Kobayashi M, et al. Modeling and cost-effectiveness analysis of etanercept in adults with rheumatoid arthritis in Japan: a preliminary analysis. Mod Rheumatol. 2006;16(2):77–84.

123. Tebas P, Henry K, Nease R, Murphy R, Phair J, Powderly WG. Timing of antiretroviral therapy. Use of Markov modeling and decision analysis to evaluate the long-term implications of therapy. Aids. 2001;15(5):591-9.

124. Tran-Duy A, Boonen A, Kievit W, van Riel PL, van de Laar MA, Severens JL. Modelling outcomes of complex treatment strategies following a clinical guideline for treatment decisions in patients with rhematoid arthritis. Pharmacoeconomics. 2014;32(10):1015–28.

125. Wailoo A, Bansback N, Brennan A, Nixon R, Michaud K, Wolfe F. Modeling the Cost Effectiveness of Etanercept, Adalimumab and Anakinra Compared to Infliximab in the Treatment of Patients with Rheumatoid Arthritis in the Medicare Program: AHRQ, U.S. Department of Health and Human Services2006.

126. Welsing PMJ, Severens JL, Hartman M, van Riel PLCM, Laan RFJM. Modeling the 5-year cost effectiveness of treatment strategies including tumor necrosis factor-blocking agents and leflunomide for treating rheumatoid arthritis in the Netherlands. Arthritis Rheumatol. 2005;51(6):964–73. https://doi.org/10.1002/art.20843.

127. Wong Y-N, Meropol NJ, Speier W, Sargent D, Goldberg RM, Beck JR. Cost implications of new treatments for advanced colorectal cancer. Cancer. 2009;115(10):2081–91. https://doi.org/10.1002/cncr.24246.

128. Woolacott N, Hawkins N, Mason A, Kainth A, Khadjesari Z, Vergel YB, et al. Etanercept and efalizumab for the treatment of psoriasis: a systematic review. Health Technol Assess. 2006;10(46):1–233.

129. Wu B, Wilson A, Wang Ff, Wang Sl, Wallace DJ, Weisman MH et al. Cost effectiveness of different treatment strategies in the treatment of patients with moderate to severe rheumatoid arthritis in China. PLoS ONE [Electronic Resource]. 2012;7(10). https://doi.org/10.1371/journal.pone.0047373.

130. Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. Health Econ. 2006;15(12):1295–310.

131. Kim SY, Goldie SJ. Cost-effectiveness analyses of vaccination programmes: a focused review of modelling approaches. Pharmacoeconomics. 2008;26(3):191–215.

132. Barton P, Bryan S, Robinson S. Modelling in the economic evaluation of health care: selecting the appropriate approach. J Health Serv Res Policy. 2004;9(2):110–8.

133. Bentley TG, Kuntz KM, Ringel JS. Bias associated with failing to incorporate dependence on event history in Markov models. Med Decis Mak. 2010. https://doi.org/10.1177/0272989X10363480.

134. Caro JJ, Briggs AH, Siebert U, Kuntz KM, ISPOR-SMDM modeling good research practices task force. Modeling good research practices—overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-1. Value in Health. 2012;15(6):796-803. https://doi.org/10.1016/j.jval.2012.06.012.

135. Caro JJ, Möller J, Getsios D. Discrete event simulation: the preferred technique for health economic evaluations? Value Health. 2010;13(8):1056–60.

136. Cooper K, Brailsford SC, Davies R. Choice of modelling technique for evaluating health care interventions. J Oper Res Soc. 2007;58(2):168–76. https://doi.org/10.1057/palgrave.jors.2602230.

137. Karnon J. Alternative decision modelling techniques for the evaluation of health care technologies: Markov processes versus discrete event simulation. Health Econ. 2003;12(10):837–48.

138. Karnon J, Haji Ali Afzali H. When to use discrete event simulation (DES) for the economic evaluation of health technologies? A review and critique of the costs and benefits of DES. Pharmacoeconomics. 2014;32(6):547–58. https://doi.org/10.1007/s40273-014-0147-9.

139. Karnon J, Stahl J, Brennan A, Caro JJ, Mar J, Möller J, et al. Modeling using discrete event simulation: a report of the ISPOR-SMDM modeling good research practices task force-4. Value Health. 2012;15(6):821–7. https://doi.org/10.1016/j.jval.2012.04.013.

140. O'Mahony JF, Newall AT, van Rosmalen J. Dealing with time in health economic evaluation: methodological issues and recommendations for practice. Pharmacoeconomics. 2015;33(12):1255–68. https://doi.org/10.1007/s40273-015-0309-4.

141. Roberts M, Russell LB, Paltiel AD, Chambers M, McEwan P, Krahn M, et al. Conceptualizing a model: a report of the ISPOR-SMDM modeling good research practices task force-2. Med Decis Mak. 2012;32(5):678–89.

142. Stahl JE. Modelling methods for pharmacoeconomics and health technology assessment: an overview and guide. Pharmacoeconomics. 2008;26(2):131–48.

143. Mauskopf J, Samuel M, McBride D, Mallya UG, Feldman SR. Treatment sequencing after failure of the first biologic in cost-effectiveness models of psoriasis: a systematic review of published models and clinical practice guidelines. Pharmacoeconomics. 2014;32(4):395–409.

144. Miller JD, Foley KA, Russell MW, MAPE. Current Challenges in Health Economic Modeling of Cancer Therapies: A Research Inquiry. Am Health Drug Benefits. 2014;7(3):153-62.

145. Brixner DI, Watkins JB. Can CER be an effective tool for change in the development and assessment of new drugs and technologies? J Managed Care Pharm. 2012;18(5 Supp A):S06-11.

146. Ericson A, Gsteiger S. Deliverable D1.6 WP1 case study review: rheumatoid arthritis. wwwimi-getrealeu/Portals/1/Documents/01%20deliverables/Deliverable%2016%20Report%20-%20Rheumatoid%20Arthritis_webversionpdf 2016. Accessed Dec 2017.

147. Tafuri G, Pagnini M, Moseley J, Massari M, Petavy F, Behring A, et al. How aligned are the perspectives of EU regulators and HTA bodies? A comparative analysis of regulatory-HTA parallel scientific advice. Br J Clin Pharmacol. 2016;82(4):965–73.

148. Ciani O, Jommi C. The role of health technology assessment bodies in shaping drug development. Drug Design Dev Therapy. 2014;8:2273–81. https://doi.org/10.2147/DDDT.S49935.

149. Lewis R, Williams N, Matar HE, Din N, Fitzsimmons D, Phillips C, et al. The effectiveness and cost effectiveness of management strategies for sciatica: systematic review and economic model. Health Technol Assess. 2011;15(39):1–578.

150. Drummond MF, Iglesias CP, Cooper NJ. Systematic reviews and economic evaluations conducted for the National Institute

for Health and Clinical Excellence in the United Kingdom: a game of two halves? Int J Technol Assess Health Care. 2008;24(2):146–50.

151. Krist AH, Wolff TA, Jonas DE, Harris RP, LeFevre ML, Kemper AR, et al. Update on the methods of the US Preventive Services Task Force: methods for understanding certainty and net benefit when making recommendations. Am J Prev Med. 2018;54(1):S11–8.

152. Wahlster P, Brereton L, Burns J, Hofmann B, Mozygemba K, Oortwijn W et al. Guidance on the integrated assessment of complex health technologies - The INTEGRATE-HTA Model [Online]. https://www.integrate-hta.eu/downloads/. 2016. Accessed 29 Aug 2017.

153. Rohwer A, Booth A, Pfadenhauer L, Brereton L, Gerhardus A, Mozygemba K et al. Guidance On The use of logic models in health technology assessments of complex interventions [Online]. https://www.integrate-hta.eu/downloads/ 2016. Accessed 29 Aug 2017.

154. Squires H, Chilcott J, Akehurst R, Burr J, Kelly MP. A framework for developing the structure of public health economic models. Value Health. 2016;19(5):588–601. https://doi.org/10.1016/j.jval.2016.02.011.

155. Tappenden P. Conceptual modelling for health economic model development In: HEDS Discussion Paper 12/05. https://eprints.whiterose.ac.uk/74464/. Accessed 12 Feb 2017. 2012.

156. Alemao E, Al MJ, Boonen AA, Stevenson MD, Verstappen SMM, Michaud K, et al. Conceptual model for the health technology assessment of current and novel interventions in rheumatoid arthritis. PLoS ONE. 2018;13(10):e0205013. https://doi.org/10.1371/journal.pone.0205013.

157. Afzali HHA, Bojke L, Karnon J. Improving decision-making processes in health: is it time for (Disease-Specific) reference models? Appl Health Econ Health Policy. 2020;18:1–4. https://doi.org/10.1007/s40258-019-00510-6.

158. Ghabri S, Stevenson M, Möller J, Caro JJ. Trusting the results of model-based economic analyses: is there a pragmatic validation solution? Pharmacoeconomics. 2019;37(1):1–6. https://doi.org/10.1007/s40273-018-0711-9.

159. Dunlop WCN, Mason N, Kenworthy J, Akehurst RL. Benefits, challenges and potential strategies of open source health economic models. Pharmacoeconomics. 2017;35(1):125–8. https://doi.org/10.1007/s40273-016-0479-8.

160. Barton P. Development of the Birmingham Rheumatoid Arthritis Model: past, present and future plans. Rheumatology (Oxford). 2011;50 (Suppl 4):iv32-iv8. doi:https://doi.org/10.1093/rheumatology/ker244.

161. Tosh J, Brennan A, Wailoo A, Bansback N. The Sheffield rheumatoid arthritis health economic model. Rheumatology. 2011;50(4):26–31.

162. Johansson EC, Hartz S, Kiri SH, Kumar G, Svedbom A. Cost-effectiveness analysis of sequential biologic therapy with ixekizumab versus secukinumab as first-line treatment of moderate-to-severe psoriasis in the UK. J Med Econ. 2018;21(8):810–20. https://doi.org/10.1080/13696998.2018.1474747.

163. Schweikert B, Malmberg C, Åkerborg Ö, Kumar G, Nott D, Kiri S, et al. Cost-effectiveness analysis of sequential biologic therapy with ixekizumab versus secukinumab in the treatment of active psoriatic arthritis with concomitant moderate-to-severe psoriasis in the UK. PharmacoEcon Open. 2020. https://doi.org/10.1007/s41669-020-00202-1.

164. Makady A, Stegenga H, Ciaglia A, Debray TP, Lees M, Happich M, et al. Practical implications of using real-world evidence (RWE) in comparative effectiveness research: learnings from IMI-GetReal. J Comp Effect Res. 2017;6(6):485–90. https://doi.org/10.2217/cer-2017-0044.

165. Labrecque JA, Swanson SA. Target trial emulation: teaching epidemiology and beyond. Eur J Epidemiol. 2017;32(6):473–5. https://doi.org/10.1007/s10654-017-0293-4.

166. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. Am J Epidemiol. 2016;183(8):758–64. https://doi.org/10.1093/aje/kwv254.

167. Cain LE, Saag MS, Petersen M, May MT, Ingle SM, Logan R, et al. Using observational data to emulate a randomized trial of dynamic treatment-switching strategies: an application to antiretroviral therapy. Int J Epidemiol. 2016;45(6):2038–49. https://doi.org/10.1093/ije/dyv295.

168. Kandiyali R, Hawton A, Cabral C, Mytton JA, Shilling V, Morris C, et al. Working with patients and members of the public: informing health economics in child health research. PharmacoEcon Open. 2019;3(2):133–41. https://doi.org/10.1007/s41669-018-0099-7.

169. Vale C, Gyte G, Humphreys R, McIlwain C, Popay J, Spera N, et al. Public involvement in systematic reviews: supplement to the briefing notes for researchers. Eastleigh: INVOLVE; 2012.

170. Hyde C, Dunn KM, Higginbottom A, Chew-Graham CA. Process and impact of patient involvement in a systematic review of shared decision making in primary care consultations. Health Expect. 2017;20(2):298–308. https://doi.org/10.1111/hex.12458.

171. Concannon TW, Grant S, Welch V, Petkovic J, Selby J, Crowe S, et al. Practical guidance for involving stakeholders in health research. J Gen Intern Med. 2019;34(3):458–63. https://doi.org/10.1007/s11606-018-4738-6.

172. Pollock A, Campbell P, Struthers C, Synnot A, Nunn J, Hill S, et al. Stakeholder involvement in systematic reviews: a scoping review. Syst Rev. 2018;7(1):208. https://doi.org/10.1186/s13643-018-0852-0.

173. Pollock A, Campbell P, Struthers C, Synnot A, Nunn J, Hill S, et al. Development of the ACTIVE framework to describe stakeholder involvement in systematic reviews. J Health Serv Res Policy. 2019;24(4):245–55. https://doi.org/10.1177/1355819619841647.

174. Concannon TW, Meissner P, Grunbaum JA, McElwee N, Guise J-M, Santa J, et al. A new taxonomy for stakeholder engagement in patient-centered outcomes research. J Gen Intern Med. 2012;27(8):985–91. https://doi.org/10.1007/s11606-012-2037-1.

175. Hudson M, Suissa S. Avoiding common pitfalls in the analysis of observational studies of new treatments for rheumatoid arthritis. Arthritis Care Res. 2010;62(6):805–10. https://doi.org/10.1002/acr.20124.

176. Schneeweiss S, Gagne JJ, Glynn RJ, Ruhl M, Rassen JA. Assessing the comparative effectiveness of newly marketed medications: methodological challenges and implications for drug development. Clin Pharmacol Ther. 2011;90(6):777–90. https://doi.org/10.1038/clpt.2011.235.

177. van Vollenhoven RF. Switching between anti-tumour necrosis factors: trying to get a handle on a complex issue. Ann Rheum Dis. 2007;66(7):849–51.

178. Emery P. Optimizing outcomes in patients with rheumatoid arthritis and an inadequate response to anti-TNF treatment. Rheumatology. 2012;51(Suppl 5):v22–30. https://doi.org/10.1093/rheumatology/kes115.

179. Hetland ML, Christensen IJ, Tarp U, Dreyer L, Hansen A, Hansen IT, et al. Direct comparison of treatment responses, remission rates, and drug adherence in patients with rheumatoid arthritis treated with adalimumab, etanercept, or infliximab: results from eight years of surveillance of clinical practice in the nationwide Danish DANBIO registry. Arthritis Rheumatoly. 2010;62(1):22–32. https://doi.org/10.1002/art.27227.

180. Rubbert-Roth A, Finckh A. Treatment options in patients with rheumatoid arthritis failing initial TNF inhibitor therapy: a

critical review. Arthritis Res Therapy. 2009;11(Suppl 1):S1. https://doi.org/10.1186/ar2666.

181. Pal SK, Vogelzang NJ. Sequential treatment strategies and combination therapy regimens in metastatic renal cell carcinoma. Clin Adv Hematol Oncol. 2013;11(3):146–55.

182. McAlister FA, Laupacis A, Wells GA, Sackett DL. Users' Guides to the Medical Literature: XIX. Applying clinical trial results B. Guidelines for determining whether a drug is exerting (more than) a class effect. JAMA. 1999;282(14):1371-7. doi:https://doi.org/10.1001/jama.282.14.1371.

183. Allegra C, Blanke C, Buyse M, Goldberg R, Grothey A, Meropol NJ, et al. End points in advanced colon cancer clinical trials: a review and proposal. J Clin Oncol. 2007;25(24):3572–5.

184. Chibaudel B, Bonnetain F, Shi Q, Buyse M, Tournigand C, Sargent DJ, et al. Alternative end points to evaluate a therapeutic strategy in advanced colorectal cancer: evaluation of progression-free survival, duration of disease control, and time to failure of strategy—an Aide et Recherche en Cancerologie Digestive Group Study. J Clin Oncol. 2011;29(31):4199–204. https://doi.org/10.1200/JCO.2011.35.5867.

185. Siebert U, Alagoz O, Bayoumi AM, Jahn B, Owens DK, Cohen DJ, et al. State-transition modeling: a report of the ISPOR-SMDM modeling good research practices task force-3. Value Health. 2012;15:812–20.

186. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. Med Decis Mak. 1993;13(4):322–38.