Integrated evaluation of targeted and non-targeted therapies in a network meta-analysis

Tanja Proctor | Katrin Jensen | Meinhard Kieser

Institute of Medical Biometry and Informatics, Heidelberg, Germany

Correspondence
Tanja Proctor, Institute of Medical Biometry and Informatics, Im Neuenheimer Feld 130.3, 69120 Heidelberg, Germany.
Email: proctor@imbi.uni-heidelberg.de

Correction added on 28 October 2020, after first online publication: Projekt Deal funding statement has been added.

Funding information
Bundesministerium für Bildung und Forschung, Grant/Award Number: 01EK1504

Abstract
Individualized therapies for patients with biomarkers are moving more and more into the focus of research interest when developing new treatments. Hereby, the term individualized (or targeted) therapy denotes a treatment specifically developed for biomarker-positive patients. A network meta-analysis model for a binary endpoint combining the evidence for a targeted therapy from individual patient data with the evidence for a non-targeted therapy from aggregate data is presented and investigated. The biomarker status of the patients is either available at patient-level in individual patient data or at study-level in aggregate data. Both types of biomarker information have to be included. The evidence synthesis model follows a Bayesian approach and applies a meta-regression to the studies with aggregate data. In a simulation study, we address three treatment arms, one of them investigating a targeted therapy. The bias and the root-mean-square error of the treatment effect estimate for the subgroup of biomarker-positive patients based on studies with aggregate data are investigated. Thereby, the meta-regression approach is compared to approaches applying alternative solutions. The regression approach has a surprisingly small bias even in the presence of few studies. By contrast, the root-mean-square error is relatively greater. An illustrative example is provided demonstrating implementation of the presented network meta-analysis model in a clinical setting.

KEYWORDS
binary endpoint, biomarker status, individual patient data, meta-regression, network meta-analysis

1 | INTRODUCTION

Systematic reviews and meta-analyses have become increasingly important in health care (Egger, Smith, & Altman, 2001; Evans, Thornton, Chalmers, & Glasziou, 2011) and can reach a high level of evidence if the included trials are of high quality (CEBM, 2018). Especially, network meta-analyses provide useful evidence for judiciously selecting the best treatment if more than one therapy option is available but each included trial only compares a subset of these treatment options (Dias, Ades, Welton, Jansen, & Sutton, 2018; Hoaglin et al., 2011). Direct and indirect evidence is combined in network meta-analysis. At present, a main focus in clinical research is set on targeted therapies in order to treat subgroups of patients taking into account their individual composition of disease, for example, the presence or absence of a special biomarker. Network meta-analysis methods need to be investigated as to whether the integrated evaluation of these new targeted therapies is feasible. The present paper investigates a...
network meta-analysis model that is able to synthesize evidence for targeted therapies on patient subgroups and for non-targeted therapies on a mixed patient population with regard to a special biomarker status, when there are no separate treatment effect estimates for biomarker-positive patients available. A synthesis model for a binary outcome is considered where individual patient data (IPD) and aggregate data (AD) are available, and information on the biomarker status of the patients should be included.

For IPD, the biomarker information is given at patient-level, for AD at study-level. Thompson and Higgins (2005) already published their thoughts on whether meta-analysis can help to target interventions at individuals most likely to benefit. Meta-regression was discussed, which allows to relate the treatment effect in every trial to some average characteristic of the patients in that trial (such as the proportion of biomarker-positive patients). Although interpretation of meta-regression results is not straightforward because of a possible ecological bias, it is the only way to analyze aggregate data taking into consideration some information about patient characteristics. Therefore, a meta-regression approach is part of the synthesis model if sufficient trials with respect to AD are available. Nevertheless, alternative approaches which work as kind of imputation methods in treating the missing biomarker as a missing data problem have to be discussed. Model bias is investigated in a simulation study. Application issues are demonstrated by a motivating clinical example.

After describing the motivating clinical example in Section 2, Section 3 outlines a model for network meta-analysis in the presence of IPD and AD. Covariates on patient- as well as study-level will be included in the model in order to incorporate information on the biomarker status for each patient or on the percentage of patients in a study with a certain biomarker status, respectively. Alternative methods without meta-regression approach are presented in Section 4. Model bias of the synthesis models are investigated by a simulation study in Section 5, and the model is applied to our motivating clinical example in Section 6. A discussion given in Section 7 ends the paper.

2 | MOTIVATING EXAMPLE

One example of individualized therapy is the treatment for nonsmall cell lung cancer (NSCLC). Here, in recent years results of studies which compared targeted agents such as Erlotinib (Rosell et al., 2012; Wu et al., 2015; Zhou, Wu, Chen, & Feng, 2011) or Gefitinib (Han et al., 2012; Kim et al., 2008; Maemondo et al., 2010; Maruyama et al., 2008; Mitsudomi et al., 2010; Mok et al., 2009; Yang et al., 2014; Zhou et al., 2014) to control treatments, that is chemotherapeutic agents as well as studies comparing Erlotinib to Gefitinib (Urata et al., 2016; Xie, Liang, & Su, 2015; Yang et al., 2017) were published. Erlotinib and Gefitinib are both receptor tyrosine kinase inhibitors, which act on the epidermal growth factor receptor (EGFR). Early studies of Gefitinib or Erlotinib included patients independent of their EGFR status (Han et al., 2012; Mok et al., 2009). EGFR belongs to a group of biomarkers. In general, there are different and sometimes inconsistent definitions of which patients are defined as biomarker-positive. These definitions depend on the respective biomarker and the method which is used to determine the biomarker status. In our motivating example, the definitions of the single studies are used and biomarker-positive and biomarker-negative patients are considered without further differentiation. In the following, we focus on binary endpoints and use as outcome of interest in the aforementioned studies the overall response rate (ORR), which is reported as a secondary endpoint in most clinical cancer studies.

The main aim of our network meta-analysis model in this context is to use all evidence available in order to conduct a network meta-analysis with targeted treatments such as Erlotinib ($E_+$), standard therapies ($S$) such as Gefitinib, and control treatments ($C$) (see Figure 1). In our clinical example, we combined the chemotherapeutical treatments with which the treatments Erlotinib and Gefitinib were compared (see Table 1) as control group. This leads to an increased precision. When considering a binary
TABLE 1 Studies used in the motivational example

| Study            | Intervention       | ORR*       | Control                          | ORR*       | Prop. BM b |
|------------------|--------------------|------------|----------------------------------|------------|------------|
| ENSURE (Wu et al., 2015) | Erlotinib ($E_+$) | 69/110     | Gemcitabine + Cisplatin ($C$)  | 36/107     | 100%       |
| EURTAC (Rosell et al., 2012) | Erlotinib ($E_+$) | 50/86      | Cisplatin + Docetaxel/Gemcitabine ($C$) | 13/87     | 100%       |
| OPTIMAL (Zhou, Wu, Chen, & Feng, 2011) | Erlotinib ($E_+$) | 68/82      | Gemcitabine + Carboplatin ($C$) | 26/72      | 100%       |
| IPASS (Mok et al., 2009) | Gefitinib ($S$)   | 95/223     | Carboplatin + Paclitaxel ($C$) | 81/214     | 59.7%b     |
| NEJGSG (Maenomdo et al., 2010) | Gefitinib ($S$)   | 84/114     | Carboplatin + Paclitaxel ($C$) | 35/114     | 100%       |
| First Signal (Han et al., 2012) | Gefitinib ($S$)   | 29/53      | Gemcitabine + Cisplatin ($C$)  | 20/42      | 44%b       |
| WJTOG (Mitsudomi et al., 2010) | Gefitinib ($S$)   | 36/58      | Cisplatin + Docetaxel ($C$)    | 19/59      | 100%       |
| Yang (Yang et al., 2014) | Gefitinib ($S$)  | 17/35      | Pemetrexed+Cisplatin ($C$)     | 20/39      | 67.7%      |
| V-15-32c (Maruyama et al., 2008) | Gefitinib ($S$) | 6/9       | Docetaxel ($C$)                 | 5/11        | 100%       |
| INTERESTc (Kim et al., 2008) | Gefitinib ($S$) | 15/125     | Docetaxel ($C$)                 | 16/142     | 14.23%     |
| Zhouc (Zhou et al., 2014) | Gefitinib ($S$)  | 6/54      | Pemetrexed ($C$)                | 6/54       | 71%        |
| Yangc (Yang et al., 2017) | Erlotinib ($E_+$) | 76/128     | Gefitinib ($S$)                | 67/128     | 100%       |
| Xiec (Xie, Liang, & Su, 2015) | Erlotinib ($E_+$) | 14/23     | Gefitinib ($S$)                | 15/27      | 100%       |
| WJOG5108Lc (Urata et al., 2016) | Erlotinib ($E_+$) | 93/201   | Gefitinib ($S$)                | 104/218    | 71.7%      |

*ORR denotes the overall response rate of the patients analyzed for biomarker status.
*bProp. BM denotes the proportion of biomarker-positive patients of the patients analyzed for biomarker status.
*cSecond line study.

endpoint and not including covariates, there is no additional benefit in using IPD instead of AD in the meta-analysis. However, the reason for considering IPD here is that this approach is meant to be extendable by including covariates or evaluating other types of endpoints like continuous or time-to-event endpoints. The data of the studies comparing $E_+$ and $C$ were implemented in the model as IPD. In the present situation, transforming the AD into IPD was straightforward because all patients are biomarker-positive and the considered outcome is the overall response rate. The information on Gefitinib studies is taken as AD. The IPASS study (Mok et al., 2009) and the WJTOG study (Mitsudomi et al., 2010) retrospectively assessed the biomarker status for some patients. We only used the data of biomarker-assessed patients in order to be able to evaluate the results of our methods in this example. The response rates and the proportion of biomarker-positive patients out of the patients analyzed for biomarker status in each study used for this motivating clinical example are listed in Table 1.

3 | SYNTHESIS MODEL FOR INDIVIDUAL PARTICIPANT AND AGGREGATE DATA INCLUDING COVARIATES

The network to be investigated is visualized in Figure 1. It consists of randomized controlled trials comparing a control treatment $C$ either with a newly developed targeted therapy $E_+$ in biomarker-positive patients or with a standard therapy $S$ in a mixed patient population of biomarker-positive and -negative patients. Whereas for the comparison of $C$ with $E_+$ individual patient data are available, only aggregate data are obtained from trial publications of the control treatment $C$ versus the standard therapy ($S$). For ease of presentation, the network model is only illustrated for the inclusion of two-arm trials. The network is assumed to be consistent, that is direct evidence agrees with indirect evidence on each treatment comparison (Salanti, 2012). The quantity we are specifically interested in is the treatment effect of the newly developed target therapy $E_+$ compared to the standard therapy $S$ in biomarker-positive patients based on all available direct and indirect evidence and, especially, based on treatment comparisons in the targeted as well as in the whole patient population. The outcome of interest is a binary outcome, for example, overall response rate.

A network meta-analysis model combining IPD and AD is required which allows to integrate treatment comparisons in patient subgroups as well as more comprehensive patient populations. In the following, a modified version of the random effects model according to Saramago, Sutton, Cooper, and Manca (2012), which follows the Sutton, Kendrick, and Coupland (2008) approach to pairwise meta-analysis, is used. It is a Bayesian arm-based (notation, see Dias and Ades, 2016) approach for a binary outcome allowing to incorporate covariates. Here, the covariate of interest is the biomarker status at patient-level in IPD and the percentage of biomarker-positive patients at study level in AD. Cluster studies are not considered in the present paper. The model includes also the approach for biomarker-negative patients $BM_-$. Therapy $E_-$ would then be a treatment only for biomarker-negative patients. In our simulations, biomarker-negative patients are not considered, but in order to keep the model as general as possible $E_-$ is included here as well.
Part I: Model for IPD studies. A modification of the random effects model proposed by Saramago et al. (2012) can be used to model the arm-based responses in the patient subgroups $BM_+$ and $BM_-$. The modification relates to the differentiation of the studies comparing $C$ and $E_+$ from the studies investigating $C$ and $E_-:$

\[
Y_{ijC} \sim \text{Bernoulli}(p_{ijC}), \ Y_{ijE_+} \sim \text{Bernoulli}(p_{ijE_+}), \ Y_{ijE_-} \sim \text{Bernoulli}(p_{ijE_-})
\]

\[
\text{logit}(p_{ijC}) = \mu_{jC}^{PD} + \delta_{jCE_+} \quad \text{with } j = 1, \ldots, \text{number of trials including } C
\]

\[
\text{logit}(p_{ijE_+}) = \mu_{jC}^{PD} + \delta_{jCE_+} \quad \text{with } j = 1, \ldots, \text{number of trials including } E_+
\]

\[
\delta_{jCE_+} \sim N\left(d_{CE_+}, \sigma^2\right), \quad \delta_{jCE_-} \sim N\left(d_{CE_-}, \sigma^2\right).
\]

where $Y_{ijC}, Y_{ijE_+}$, and $Y_{ijE_-}$ denote the binary responses of the $i$-th participant in treatment arm $C$, $E_+$, or $E_-$, respectively, of the $j$-th study (i.e., $1 = \text{event}, 0 = \text{no event}$) and are assumed to follow a Bernoulli distribution with the event occurring with probability $p_{ijC}, p_{ijE_+},$ or $p_{ijE_-}$. A standard logistic regression model is fitted to each participant $i$ of the $j$-th study, with $\mu_{jC}^{PD}$ representing the log odds for the control group $C$ in study $j$. For each study $j$, $\delta_{jCE_+}$ is the log odds ratio for treatment $E_+$ relative to the control $C$ and the analogue notation is used for $\delta_{jCE_-}$. Prior distributions need to be specified for $\mu_{jC}^{PD}, d_{CE_+}, d_{CE_-},$ and $\sigma$.

Part II: Model for AD studies including covariates. Following the notation of Saramago et al. (2012), a meta-regression random effects model for binary outcome data can be used to model the arm-based responses in dependence of the proportion of biomarker-positive patients in each study:

\[
r_{jC} \sim \text{Bin}(p_{jC}, n_{jC}), \ r_{jS} \sim \text{Bin}(p_{jS}, n_{jS})
\]

\[
\text{logit}(p_{jC}) = \mu_{jC}^{AD} \quad \text{with } j = 1, \ldots, \text{number of trials including } C
\]

\[
\text{logit}(p_{jS}) = \mu_{jC}^{AD} + \delta_{jCS} + \beta_{CS}^{B} \cdot X_j
\]

\[
\quad \text{with } j = 1, \ldots, \text{number of trials including } S
\]

\[
\delta_{jCS} \sim N\left(d_{CS}, \sigma^2\right)
\]

\[
\beta_{CS}^{B} \cdot X_j \quad \text{study-level specific covariate regression term},
\]

where $r_{jC}$ and $r_{jS}$ denote the number of observed events, and $n_{jC}$ and $n_{jS}$ the total number of individuals in treatment arm $C$ or $S$, respectively, of the $j$-th study. The probabilities of an event for the control and the standard treatment arm in each study are represented by $p_{jC}$ and $p_{jS}$. $\mu_{jC}^{AD}$ represents the log odds of an event for the control $C$ in study $j$, and $d_{CS}$ is the log odds ratio for the standard treatment $S$ relative to the control $C$. $\delta_{jCS}$, the log odds ratio for the standard $S$ relative to the control $C$ in study $j$, is assumed to be normally distributed with mean $d_{CS}$ and variance $\sigma^2$. The term $\beta_{CS}^{B} \cdot X_j$ represents a study-level specific covariate regression term for standard $S$ relative to the control $C$ for each study $j$. The index $B$ indicates the between-study relationship of the regression coefficient $\beta_{CS}^{B}$. The covariate $X_j$ represents the mean percentage of biomarker-positive patients of study $j$. Prior distributions need to be specified for $\mu_{jC}^{AD}, d_{CS},$ and $\sigma$. To incorporate trials with three or more arms, the use of the multivariate normal distribution for the intervention effects becomes necessary.

Part III: Combination of estimates of intervention effect including covariates. In analogy with Saramago et al. (2012), estimates of intervention effect for the standard treatment $S$ and for the new targeted therapy $E_+$ ($E_-$) versus the control treatment $C$ are combined to get an indirect intervention effect estimate for the standard treatment $S$ versus the new targeted therapy $E_+$ ($E_-$). For this indirect intervention effect, we assumed consistency of the network:

\[
\delta_{jE_+,S} \sim N\left(d_{E_+,S}, \sigma^2\right) \sim N\left(d_{CS} - d_{CE_+} + \beta_{CS}^{B} \cdot X_j, \sigma^2\right).
\]

\[
\delta_{jE_-,S} \sim N\left(d_{E_-,S}, \sigma^2\right) \sim N\left(d_{CS} - d_{CE_-} + \beta_{CS}^{B} \cdot X_j, \sigma^2\right).
\]

Prior distributions need to be specified for $\sigma$, and $\beta_{CS}^{B}$ is estimated in model Part II. It has to be noted that Part I to Part III can be extended to include multiarm trials by the use of the multivariate normal distribution for the intervention effects and to
address additional covariates (Saramago et al., 2012; Sutton et al., 2008). For simplification of the presentation, randomized controlled trials (RCTs) comparing directly the targeted therapy \( E_+ \) with the standard therapy \( S \) in biomarker-positive patients were not included in the model. However, they can be integrated in this synthesis model. Therefore, Part I has to be extended including IPD for the comparison of \( E_+ \) versus \( S \). IPD and AD have then to be synthesized according to Saramago et al. (2012). The synthesis model described to integrate non-targeted therapies in the evaluation of targeted therapies in a network meta-analysis must be based on the assumption of transitivity (Salanti, 2012). Transitivity is not easy to argue here because treatment \( C \) and \( S \) can be applied to biomarker-positive as well as biomarker-negative patients but \( E_+ \) and \( E_- \) only to biomarker-positive and biomarker-negative patients, respectively. Therefore, the treatment arms \( E_+ \) and \( E_- \) are not missing at random and, most important, the biomarker status is taken into consideration as effect modifier in the meta-regression Part II.

### 4 Synthesis Model With Alternative Approaches

The synthesis model presented in Section 3 consists of three parts. Part I and Part III can be applied without further adjustments, but Part II of the synthesis model comprises a meta-regression approach and has to be discussed. The Cochrane Handbook Chapter 9.6.5.5 (Higgins & Green, 2011) puts a warning that meta-regression could introduce ecological bias when conducting the regression with a covariate at patient-level data. The biomarker status of a patient is primarily collected at patient-level. Therefore, in the regression approach there might be the risk for ecological bias using the percentage of biomarker-positive patients at study-level. Furthermore, according to the Cochrane Handbook Chapter 9.6.4 (Higgins & Green, 2011) a “meta-regression should generally not be considered when there are fewer than 10 studies in a meta-analysis.” Borenstein, Hedges, Higgins, and Rothstein (2009a) do not give an explicit number but only states that in meta-analysis “an appropriately large ratio of studies to covariates” is needed. Therefore, alternative approaches to the meta-regression Part II are introduced in the following in order to take the aforementioned reasons into account.

These approaches treat the missing information on the biomarker status as a missing data problem and impute the biomarker status by allocating the response rates in different ways to the biomarker-positive patients. As in model Part II, the listed approaches assume that only the treatment effect for the entire study population as well as the percentage of biomarker-positive patients is known but not the treatment effect for the subgroup of biomarker-positive patients. With these methods, the response rates of the subgroup of biomarker-positive patients in both treatment groups are determined. Then the log odds ratio of the two treatment groups is estimated with a modified version of the model described in Section 3.

Five different scenarios for a synthesis model without meta-regression approach are described:

1. **Best-case scenario**: In this scenario, the maximum possible number of biomarker-positive patients are assigned having success in the control treatment group \( C \) as well as in the standard group \( S \).
2. **Best-case scenario only for biomarker-positive patients in the standard treatment group**: The best-case scenario is only applied to the biomarker-positive patients in the treatment group \( S \). For the biomarker-positive patients in the control group \( C \), it is assumed that the response rate is equal to the response rate of the overall population in the control group, with other words, there is no difference in the performance between the biomarker-positive subgroup and the biomarker-negative subgroup in the control group.
3. **Worst-case scenario**: In contrast to the best-case scenario, the worst-case scenario assumes that the minimum possible number of biomarker-positive patients show success in the both treatment groups \( C \) and \( S \).
4. **Worst-case scenario only for biomarker-positive patients in the standard treatment group**: In this scenario, only the biomarker-positive patients in the treatment group \( S \) perform the worst possible, whereas biomarker-positive patients in the control group \( C \) perform equivalent to the overall population. With other words, biomarker-positive patients in the control group have the same response rate as the overall population in the control group.
5. **Biomarker status not predictive for response rate**: This scenario assumes that the biomarker status of patients is not predictive for the treatment effect, that is the overall response rate in the subgroup of the biomarker-positive patients is the same as the overall response rate of the overall population for both treatment groups.

The aim of these alternative scenarios to the meta-regression approach is to define the limits of the best and worst possible treatment effect of the biomarker-positive subgroup and to compare the performance of these simple approaches with the meta-regression approach, especially when the number of studies in a meta-analysis is smaller than 10.
5 | SIMULATION STUDY

5.1 | Evaluation of results

The main interest of the simulation study was to determine which parameter settings in combination with which approaches delivered the best estimate of the treatment effect for the biomarker-positive subgroup. Thereby, we considered the situation that for the comparison $S$ versus $C$ there is only information as aggregate data available given by the percentage of biomarker-positive patients. The simulation delivered estimates of regression parameter $\beta$, and the treatment effect $d_{CS}$ for biomarker-negative patients “intercept” in different settings. The estimates for the treatment effect $d_{CS}$ for biomarker-positive patients where then calculated as the sum of the regression parameter and the intercept. These estimates were compared to the treatment effect for only biomarker-positive patients. For this comparison, the mean bias and the root-mean-square error (RMSE) of the estimated treatment effect were calculated.

5.2 | Setting

For the simulation of the studies comparing $E_+$ versus $S$ we modified the data generating model (DGM) “Modification of DGM ‘modified’” presented by Seide, Jensen, and Kieser (2019). This approach can be used for simulating data in the setting of multiarm trials in a random-effects network meta-analysis. In our case, we assumed a three-arm study, where one treatment arm only includes biomarker-positive patients, where the second arm considers the same treatment but only includes biomarker-negative patients, and where the control arm includes both populations. The simulation delivered for given treatment effects, covariate ranges, number of studies, and heterogeneities binary response data for biomarker-positive and biomarker-negative patients in the treatment group as well as in the control group. The response rates for biomarker-positive patients in the treatment and control group were taken in the network meta-analysis model to calculate the true treatment effects. The combined outcomes of the biomarker-positive and biomarker negative patients as well as the covariate information were used in the network meta-analysis model for evaluating the different scenarios. In all simulation scenarios, a network meta-analysis as introduced in Section 3 using a modified code from Saramago et al. (2012) was performed with the meta-regression in Part II of the model. For the alternative approaches described in Section 4, the Part II of the model introduced in Section 3 was simplified to a meta-analysis without regression. Instead, the numbers of positive responses (success) in each biomarker subgroup were calculated according to the assumptions explained in Section 4. In all scenarios of the simulation, the response rates comparing $E_+$ versus $C$ were taken from the studies displayed in Table 1 in Section 2 and were assumed to be available in IPD. The response rate used for the study comparing $E_+$ versus $S$ was constructed in a way that consistency of the network meta-analysis was assured. Seven hundred sixty-eight scenarios were evaluated for the comparison of treatment $S$ to treatment $C$. The scenarios for generating the data were all based on binary outcome data, whereby the assumed true response rates for the control group $C$ and the biomarker-negative patients in the standard treatment group $S$ stayed fixed in all scenarios with 0.4 and 0.45, respectively. This resulted in treatment effects given as log odds ratio for biomarker-negative patients as $\theta_{neg} = 0.2$. The assumed true response rate for biomarker-positive patients in the treatment group $S$ varied $(0.35, 0.45, 0.5, 0.7)$, which resulted in varying treatment effects $(\theta = (-0.21, 0.2, 0.4, 1.25))$, respectively. The number of trials comparing $S$ versus $C$ was varied as $n = (2, 5, 10, 20)$ and the between-trial heterogeneity by using values $\tau = (0.001, 0.5, 1, 2)$. The sample size of $n = 200$ was assumed to be the same in all trials of the same meta-analysis and equal between treatment groups $S$ and control $C$. The two different ranges $cov = (0.3 – 0.7; 0.4 – 0.5)$ which were used for the covariate, that is the percentage of biomarker-positive patients, were divided in equidistant steps depending on the number of studies per meta-analysis. As in Saramago et al. (2012), we used vague priors with $\delta \sim N(0, 10^6)$ and $\sigma \sim Unif(0, 2)$. Each of the scenarios was simulated a 1,000 times. All simulations were performed in $\text{R}$ version 3.2.4 ($\text{R}$ Core Team, 2016). The $\text{R}$ package $\text{R2jags}$ (Su & Yajima, 2015) was used in order to conduct the Bayesian analysis in $\text{JAGS}$ (Plummer, 2003). The number of burn-ins in $\text{JAGS}$ was set to 20,000 with 50,000 iterations.

5.3 | Results

The results of the simulation study are displayed in the trellis plots in Figures 2 and 3. R-code to reproduce the results of the simulation study is available as Supporting Information on the journals web page (http://onlinelibrary.wiley.com/doi/10.1002/bimj.201800322/suppinfo). In our simulation study, we always conducted a network meta-analysis, but solely changed the settings of Part II of the model, that is the comparison of $S$ versus $C$. Therefore, we evaluated the treatment effect estimate $S$ versus $C$. 

**FIGURE 2** Mean bias (A) and RMSE (root mean square error) (B) of the network meta-analysis with the meta-regression approach for the treatment effect estimate of $S$ compared to $C$ for biomarker-positive patients are displayed with varying assumed true treatment effects $\theta$, number of studies per meta-analysis, covariate ranges, and heterogeneities $\tau$.

**Meta-regression approach**

Figure 2 displays the mean bias (on top) as well as the RMSE (below) in estimating the treatment effect for the biomarker-positive patients subgroup for the standard treatment $S$ compared to the control treatment $C$ with a network meta-analysis including the meta-regression approach applied in Part II of the model for 2 - 20 studies with the treatment effects of the biomarker-positive patients given as log odds ratios $\theta = (-0.21, 0.2, 0.4, 1.25)$ and the covariate ranges $cov = 0.3 - 0.7$ and $cov = 0.4 - 0.5$. 
Mean bias and RMSE (root mean squared error) of the TE estimate of $S$ compared to $C$ for biomarker-positive patients are displayed for varying methods, assumed true treatment effects $\theta$, number of studies, covariate ranges, and heterogeneities $\tau$.

**Bias**

As can be seen in Figure 2 (on top) the extent of the bias is influenced by the number of studies especially for a strong treatment effect $\theta = (-0.21, 1.25)$ and a low heterogeneity $\tau = (0.001, 0.5)$. The regression seems to overestimate the treatment effect when the true treatment effect for
biomarker-positive patients ($\theta = -0.21$) is smaller than the treatment effect for biomarker-negative patients ($\theta = 0.2$) and underestimates the treatment effect when the true treatment effect for biomarker-positive patients is higher ($\theta = -0.21$).

We also noticed that the covariate range influences the bias. For smaller covariate range the bias is increasing, which might be due to the fact that with a smaller range of covariates the regression becomes more unstable and thus the extrapolation for the upper limit, a covariate value of one, gets more unstable. With a higher heterogeneity, especially for $\tau = 2$ the bias increases and the estimates are less reliable.

**RMSE**

The RMSE of the estimated treatment effect, displayed in Figure 2 (bottom) gives more weight to larger deviations in the different simulation runs compared to the bias.

A covariate range between 0.4 and 0.5 combined with high treatment effects leads to an increasing RMSE when the number of studies is increasing. For other values of $\theta$ it can be noticed that the RMSE is always higher for smaller covariate ranges compared to larger covariate ranges which we expected because variance increases with a smaller range of independent values and is inversely related to the spread of covariate variables. The higher the heterogeneity the higher the RMSE becomes.

**Alternative approaches**

Figure 3 illustrates the mean bias (on top) and the RMSE (below) of the estimated treatment effect, given as log odds ratio, when comparing $S$ to $C$ for the biomarker-positive patient subgroup for all methods evaluated in the simulation study. The different scenarios are marked with different shapes, and the two covariate ranges are given in black and gray. The estimation based on the “best-case only for biomarker-positive patients in the standard treatment group ($S_2$)” scenario produced larger absolute bias and RMSEs under all simulation settings compared to the “best-case” scenario. In this scenario, the biomarker-positive patients in the treatment group $S$ are assumed to have a considerably increased response rate compared to the biomarker-positive patients in the control group whereas the assumed true treatment effect of biomarker-positive patients is not as large. If we look at the logit curve, the points are also at the tails of distribution, and this might be as well a reason for this notable bias. The estimate based on the scenario “worst-case only for biomarker-positive patients in the treatment group ($S_1$)” leads to larger absolute bias and a larger RMSE throughout all settings compared to the “worst-case” scenario, because it always underestimates the assumed true treatment. For ease of presentation these scenarios are not displayed in Figure 3 but only in the Supplement in Figure S1.

**Bias**

As can be seen in Figure 3, the treatment effect estimation with the meta-regression approach (diamond) produced throughout all settings the smallest bias compared to the other scenarios. For $\theta = (-0.21, 0.2, 0.4)$ the estimation using the approach “biomarker-positive patients are not predictive” (star) generated an equal bias compared to the results of the meta-regression approach and a smaller bias than the other alternative approaches and slightly underestimates the treatment effect for $\theta = 1.25$. This result is not surprising because it is assumed that the proportion of patients with positive response rate in the biomarker-positive subgroup is equal to those of the biomarker-negative subgroup which again is equal to the proportion of patients in the overall population.

For $\theta = (1.25, 0.4, 0.2)$, the treatment effect estimation with the “best-case” scenario (the maximum number possible of biomarker patients are assigned having success) (square) leads to an overestimation of the treatment effect and therefore a negative bias. The bias decreases with an increasing number of studies. The estimation with the “best-case” scenario leads to the smallest bias for $\theta = -0.21$ and $\tau = (0.001, 0.5)$. The “worst-case” scenario (circle) produces throughout all scenarios a comparatively small bias for a covariate range of 0.3 – 0.7. Throughout all settings and all scenarios, we can see an increase in the bias for a narrower covariate range, that is for the covariate range 0.3 – 0.7 (black) compared to the covariate range 0.4 – 0.5 (gray).

In general, it can be observed that the mean bias increases with an increasing heterogeneity and decreases with an increasing number of studies.

**RMSE**

Throughout all settings the RMSE increases with increasing heterogeneity. The scenario “biomarker not predictive” produces overall the smallest RMSE. The estimates created using meta-regression approach were throughout all settings smaller than the “best-case” and “worst-case” approach. The estimation using the “best-case” approach resulted in low RMSEs for $\theta = -0.21$ and a covariate range of 0.3 – 0.7. The highest RMSEs are produced for $n < 10$, especially with $\theta = 1.25$. The “worst-case” scenario
Methods

FIGURE 4 Treatment effect estimate and credible interval of Erlotinib compared Gefitinib for biomarker-positive patients

is throughout all settings lower or equal to the “best-case” scenario. It leads to the smallest RMSEs for $\theta = (-0.21, 0.2, 0.4)$, $\tau = 0.001$ and a covariate range $0.3-0.7$.

6 | APPLICATION

We used the models introduced in Section 3 and the scenarios of Section 4 to conduct a network meta-analysis of the studies displayed in Section 2. The plain outcome numbers calculated for the alternative scenarios are given in the Supplement (Table S1). The different approaches used in Part II of the model did not influence the estimate of the treatment effect of Erlotinib versus a control treatment $C$. Therefore, the differences which are observed in Figure 4 mainly result from the differences in estimates of the treatment effect for the standard treatment Gefitinib $S$ compared with control treatment $C$. R-code to reproduce the results of the application in the motivational example is available as Supporting Information on the journals web page (http://onlinelibrary.wiley.com/doi/xxx/supplinfo). Figure 4 shows that all credible intervals except for the “worst-case approach” (filled circle) include the estimate of the true treatment effect (star). The approach “worst-case” for biomarker-positive patients in both treatments underestimates the treatment effect of Gefitinib ($S$). Gefitinib belongs as well to the tyrosine kinase inhibitors and is therefore likely to have a higher treatment effect on biomarker-positive patients. Scenario “worst-case only for $S_+$” delivered the highest odds ratio, because it underestimated the treatment effect of Gefitinib greatly and therefore overestimated the treatment effect of Erlotinib. It also resulted in the widest credible interval. The treatment effect estimate created by the meta-regression approach overestimates the true treatment effect, but it still includes it in the credible interval. The credible interval of the treatment effect estimate with meta-regression approach is quite wide, despite the fact that eight studies were included in the regression. One reason for this may be that, out of the eight studies, two were really small and that the biomarker range was not well balanced. Applying the “best-case approach only for biomarker-positive patients” (empty square) results in an overestimation of the treatment effect for biomarker-positive patients in the Gefitinib group and therefore an underestimation of the true treatment effect of Gefitinib compared to Erlotinib. Scenario “biomarker not predictive” underestimates the treatment effect of Gefitinib compared to the control group and therefore overestimates the treatment effect of Erlotinib compared to Gefitinib. In this motivational example, different chemotherapeutic agents were pooled. This may mask heterogeneity and challenge the transitivity assumption. Therefore, the results have to be interpreted with care. We did not perform additional subgroup analyses.
on the different chemotherapeutic agents, because the main aim here was to illustrate how the model works with real data and not to give therapeutic recommendations based on pooled treatment effects. However, we conducted a subgroup analysis of only first-line studies. The results are displayed in the Supplement in Figure S2.

7 | DISCUSSION

The aim of this paper was to estimate the treatment effect for a patient subgroup in a network meta-analysis based on direct and indirect evidence integrating comparisons of targeted as well as non-targeted therapies. Evidence on non-targeted therapies was characterized here by the response rates for an overall population including biomarker-positive and biomarker-negative patients and the percentage of biomarker-positive patients. The network meta-analysis model used involved a meta-regression (Part II of the model addressing aggregate data) which was investigated in detail by a simulation study. The advantage of a meta-regression is that no assumption of the treatment performance on biomarker-positive or biomarker-negative patients is needed. However, a meta-regression with only a few studies shows a very high variance and should not be considered according to the Cochrane Handbook 9.6.4 (Higgins & Green, 2011) when the number of studies included is smaller than 10. Another limitation of using the meta-regression approach based at study-level data instead of patient-level covariate data is the risk of ecological bias. By including also studies with biomarker-positive patients only or retrieving some IPD, this bias may be decreased. When concentrating on only one covariate as effect modifier, one has to be aware that the problem of confounding may arise and therefore results have to be interpreted with care (Lipsey, 2003). According to Borenstein, Hedges, Higgins, and Rothstein (2009b), small studies, few studies, or a small covariate effect could lead to problems with low power in meta-regression. Furthermore regressing on characteristics which do not differentiate substantially across trials may lead to higher biases, as could be seen in the results of the simulation study for the covariate range 0.4 - 0.5. The meta-regression model is regarded as starting point on how to use all evidence available. The aforementioned limitations show that further research in this field is needed. Alternative approaches to meta-regression were considered as well in the simulation study. In summary, the meta-regression approach had the lowest bias in comparison to all considered alternative approaches even when the number of studies were small. Nonetheless, the variance can be quite high and in consequence the credible intervals for the treatment effect estimate wide when the meta-regression approach is used for less than 10 studies or small studies. The results of the simulation study suggest that when only few studies provide aggregate data and there is high heterogeneity the approach “biomarker not predictive for response rate” is the recommended method. This approach seems to always lead to a more conservative estimate of the treatment effect but is especially appropriate if there is a high percentage of biomarker-positive patients in the overall patient population. If there are reasonable grounds to believe that the biomarker-positive subgroup performs better than the overall population then a “best-case” analysis is recommended. In the application the approaches “best-case only for S₁₁,” “best-case,” and the meta-regression performed quite well. Therefore, we suggest to use these three approaches in order to determine the possible limits of the treatment effect. The different approaches are limited by possibly to crude assumptions, that is, of the best-case and worst-case approaches. The assumption that the percentage of biomarker-positive patients is the same in both treatment groups implies a perfectly balanced randomization, which might be another limiting factor. Additionally, it could be the case that the percentage of biomarker-positive patients is not known, for example, when studies are conducted without considering the biomarker status in the analyses. On the other side, however, if biomarker tests are conducted, then often subgroup analyses are performed (see Mok et al., 2009; Han et al., 2012) and therefore none of the approaches discussed are needed in order to perform a network meta-analysis for the treatment effect of biomarker-positive patients.

We assumed a consistent network. The results might differ when applying the approaches to an inconsistent network. Vague priors were used for the Bayesian analysis in the simulation study. There was not enough knowledge in order to use informative priors, and it would have been beyond the scope of this article to investigate the influence of multiple different priors.

The transitivity assumption can be challenged in these approaches. Further research is needed on how to address this issue adequately. Recent work in Ishak, Proskorovsky, & Benedict (2015) used the matching-adjusted indirect comparison (MAIC) and simulated treatment comparison (STC) approaches to address a similar problem. These approaches would not work in our situation, because we consider 100% biomarker-positive patients in the IPD and we cannot adjust the proportion of biomarker patients to the percentage of aggregate data, which is less than 100%. Nevertheless these approaches show as well that meta-analyses for heterogeneous populations are a required field of research. In further research, we would like to extend the model for continuous and time-to-event data and might also include three-armed trials.

ACKNOWLEDGEMENTS

This study was funded by the Federal Ministry of Education and Research (BMBF) (01EK1504).
We also thank the Group HTA Statistics from F. Hoffmann La Roche AG, Basel, Switzerland (particularly Pierre Ducournau and Maximo Carreras) for the valuable discussion when preparing this work. Furthermore, we thank two anonymous reviewers and the associate editor for their critical comments, which helped to improve the presentation.

Open access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

The authors have declared no conflict of interest.

ORCID

Tanja Proctor https://orcid.org/0000-0003-3760-9971
Katrin Jensen https://orcid.org/0000-0002-5088-5279
Meinhard Kieser https://orcid.org/0000-0003-2402-4333

REFERENCES

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009a). Meta-regression. In Introduction to meta-analysis (Chap. 20, pp. 187–203). Hoboken, NJ: Wiley.
Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009b). Notes on subgroup analyses and meta-regression. In Introduction to meta-analysis (Chap. 21, pp. 205–212). Hoboken, NJ: Wiley.
CEBM (2018). Centre for evidence based medicine. Retrieved from http://www.cebm.net. Accessed: September 3, 2018.
Dias, S., & Ades, A. E. (2016). Absolute or relative effects? Arm-based synthesis of trial data. Research Synthesis Methods, 7, 23–28.
Dias, S., Ades, A. E., Welton, N., Jansen, J. P., & Sutton, A. (2018). Network meta-analysis for decision-making, Volume 1. Hoboken, NJ: Wiley.
Egger, M., Smith, G., & Altman, D. (Eds.). (2001). Systematic reviews in health care: Meta-analysis in context. London: BMJ Publishing Group.
Evans, I., Thornton, H., Chalmers, I., & Glasziou, P. (2011). Testing treatments: Better research for better healthcare (2nd ed.). London: Pinter and Martin. Retrieved from www.testingtreatments.org.
Han, J.-Y., Park, K., Kim, S.-W., Lee, D. H., Kim, H. Y., Kim, H. T., … Lee, J. S. (2012). First-SIGNAL: First-line single-agent Iressa versus gemcitabine and cisplatin trial in never-smokers with adenocarcinoma of the lung. Journal of Clinical Oncology, 30, 1122–1128.
Higgins, J., & Green, S. (2011). Cochrane-handbook for systematic reviews of interventions, version 5.1.0 [updated March 2011]. London: The Cochrane Collaboration. Retrieved from www.cochrane.org.
Hoaglin, D., Hawkins, N., Jansen, J., Scott, A., Itzler, R., Cappelleri, J., … Barrett, A. (2011). Conducting indirect-treatment-comparison and network-meta-analysis studies: Report of the ISPOR task force on indirect treatment comparisons good research practices part II. Value in Health, 14, 429–437.
Ishak, J., Proskorovsky, I., & Benedict, A. (2015). Simulation and matching-based approaches for indirect comparison of treatments. PharmacoEconomics, 33, 537–549.
Kim, E. S., Hirsh, V., Mok, T., Socinski, M. A., Gervais, R., Wu, Y.-L., … Douillard, J.-Y. (2008). Gefitinib versus docetaxel in previously treated non-small-cell lung cancer (interest): A randomised phase III trial. The Lancet, 372(9652), 1809–1818.
Lipsey, M. W. (2003). Those confounded moderators in meta-analysis: good, bad, and ugly. The Annals of the American Academy of Political and Social Science, 587(1), 69–81.
Maemondo, M., Inoue, A., Kobayashi, K., Sugawara, S., Oizumi, S., Isobe, H., … Nukiwa, T. (2010). Gefitinib or chemotherapy for nonsmall-cell lung cancer with mutated EGFR. New England Journal of Medicine, 362, 2380–2388.
Maruyama, R., Nishiwaki, Y., Tamura, T., Yamamoto, N., Tsuboi, M., Nakagawa, K., … Fukuoka, M. (2008). Phase III study, v-15-32, of gefitinib versus docetaxel in previously treated Japanese patients with nonsmall-cell lung cancer. Journal of Clinical Oncology, 26(26), 4244–4252.
Mitsudomi, T., Morita, S., Yatabe, Y., Negoro, S., Okamoto, I., Tsurutani, J., … Fukuoka, M. (2010). Gefitinib versus cisplatin plus docetaxel in patients with non-small-cell lung cancer harbouring mutations of the epidermal growth factor receptor (WJTOG3405): An open label, randomised phase 3 trial. The Lancet Oncology, 11, 121–128.
Mok, T. S., Wu, Y.-L., Thongprasert, S., Yang, C.-H., Chu, D.-T., Saijo, N., … Fukuoka, M. (2009). Gefitinib or Carboplatin-Paclitaxel in Pulmonary Adenocarcinoma. New England Journal of Medicine, 361, 947–957.
Plummer, M. (2003). Jags: A program for analysis of Bayesian graphical models using Gibbs sampling. Paper presented at Third International Workshop on “Distributed Statistical Computing” (DSC 2003), Vienna, Austria. Retrieved from https://www.r-project.org/conferences/DSC-2003-Proceedings/Plummer.pdf.
R Core Team (2016). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.
Rosell, R., Carcereny, E., Gervais, R., Vergnenegre, A., Massuti, B., Felip, E., … Paz-Ares, L. (2012). Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced (EGFR) mutation-positive non-small-cell lung cancer (EURTAC): A multicentre, open-label, randomised phase 3 trial. The Lancet Oncology, 13, 239–246.
Salanti, G. (2012). Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: Many names, many benefits, many concerns for the next generation evidence synthesis tool. Research Synthesis Methods, 3, 80–97.
Saramago, P., Sutton, A. J., Cooper, N. J., & Manca, A. (2012). Mixed treatment comparisons using aggregate and individual participant level data. *Statistics in Medicine, 31*, 3516–3536.

Seide, S., Jensen, K., & Kieser, M. (2019). Simulation and data-generation for random-effects network meta-analysis of binary outcome. *Statistics in Medicine, 38*, 3288–3303.

Su, Y.-S., & Yajima, M. (2015). *R2jags: Using R to Run 'JAGS*. R package version 0.5-7. https://CRAN.R-project.org/package=R2jags

Sutton, A. J., Kendrick, D., & Coupland, C. A. C. (2008). Meta-analysis of individual- and aggregate-level data. *Statistics in Medicine, 27*, 651–669.

Thompson, S., & Higgins, J. (2005). Can meta-analysis help target interventions at individuals most likely to benefit? *The Lancet, 365*, 341–346.

Urata, Y., Katakami, N., Morita, S., Kaji, R., Yoshioka, H., Seto, T., … Nakanishi, Y. (2016). Randomized phase III study comparing gefitinib with erlotinib in patients with previously treated advanced lung adenocarcinoma: Wjog 5108l. *Journal of Clinical Oncology, 34*(27), 3248–3257.

Wu, Y.-L., Zhou, C., Liam, C.-K., Wu, G., Liu, X., Zhong, Z., … Zuo, Y. (2015). First-line erlotinib versus gemcitabine/cisplatin in patients with advanced EGFR mutation-positive non-small-cell lung cancer: analyses from the phase III, randomized, open-label, ENSURE study. *Annals of Oncology, 26*, 1883.

Xie, Y., Liang, J., & Su, N. (2015). Gefitinib versus erlotinib as first-line treatment for patients with advanced EGFR mutation-positive non-small-cell lung cancer. *Nan Fang Yi Ke Da Xue Xue Bao, 35*, 446–449.

Yang, J. C.-H., Kang, J. H., Mok, T., Ahn, M.-J., Srimuninnimit, V., Lin, C.-C., … Park, K. (2014). First-line pemetrexed plus cisplatin followed by gefitinib maintenance therapy versus gefitinib monotherapy in east Asian patients with locally advanced or metastatic non-squamous non-small cell lung cancer: A randomised, phase 3 trial. *European Journal of Cancer, 50*(13), 2219–2230.

Yang, J.-J., Zhou, Q., Yan, H.-H., Zhang, X.-C., Chen, H.-J., Tu, H.-Y., … Wu, Y.-L. (2017). A phase III randomised controlled trial of erlotinib vs gefitinib in advanced non-small cell lung cancer with EGFR mutations. *British Journal of Cancer, 116*, 568–574.

Zhou, C., Wu, Y.-L., Chen, G., & Feng, J. E. (2011). Erlotinib versus chemotherapy as first-line treatment for patients with advanced EGFR mutation-positive non-small-cell lung cancer (OPTIMAL, CTONG-0802): A multicentre, open-label, randomised, phase 3 study. *Lancet Oncology, 12*, 735–742.

Zhou, Q., Cheng, Y., Yang, J.-J., Zhao, M.-F., Zhang, L., Zhang, X.-C., … Wu, Y.-L. (2014). Pemetrexed versus gefitinib as a second-line treatment in advanced nonsquamous nonsmall-cell lung cancer patients harboring wild-type EGFR (CTONG0806): A multicenter randomized trial. *Annals of Oncology, 25*(12), 2385–2391.

**SUPPORTING INFORMATION**

Additional supporting information including source code to reproduce the results may be found online in the Supporting Information section at the end of the article.

---

**How to cite this article:** Proctor T, Jensen K, Kieser M. Integrated evaluation of targeted and non-targeted therapies in a network meta-analysis. *Biometrical Journal*. 2020;62:777–789. https://doi.org/10.1002/bimj.201800322