Nagels HE, Rishworth JR, Siristatidis CS, Kroon B. Androgens (dehydroepiandrosterone or testosterone) for women undergoing assisted reproduction. *Cochrane Database Syst Rev* 2015;**11**:CD009749.

Nyboe Andersen A, Nelson SM, Fauser BCJM, García-Velasco JA, Klein BM, Arce J-C, ESTHER-1 study group. Individualized versus conventional ovarian stimulation for in vitro fertilization: a multicenter, randomized, controlled, assessor-blinded, phase 3 noninferiority trial. *Fertil Steril* 2017;**107**:387–396.

Oudshoorn SC, van Tilborg TC, Eijkemans MJC, Oosterhuis GJE, Friederich J, van Hooff MHA, van Santbrink EJP, Brinkhuis EA, Smeenk JMJ, Kwee J, de Koning CH *et al*. Individualized versus standard FSH dosing in women starting IVF/ICSI: an RCT. Part 2: The predicted hyper responder. *Hum Reprod* 2017;**32**:2506–2514.

Pandian Z, McTavish AR, Aucott L, Hamilton MP, Bhattacharya S. Interventions for 'poor responders' to controlled ovarian hyper stimulation (COH) in in-vitro fertilisation (IVF). *Cochrane Database Syst Rev* 2010:CD004379.

Scheinhardt MO, Lerman T, König IR, Griesinger G. Performance of prognostic modelling of high and low ovarian response to ovarian stimulation for IVF. *Hum Reprod* 2018. doi:10.1093/humrep/dey236. [Epub ahead of print].

H.L. Torrance[1,*], F.J.M. Broekmans[1], and B.W.J. Mol[2]
[1]*Department of Reproductive Medicine and Gynaecology, University Medical Centre Utrecht, Utrecht University, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands*
[2]*Department of Obstetrics and Gynaecology, Monash University, Scenic Blvd & Wellington Road, Clayton VIC 3800, Australia*

*Correspondence address. Department of Reproductive Medicine and Gynaecology, University Medical Centre Utrecht, Utrecht University, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands.
E-mail: h.torrance@umcutrecht.nl

doi:10.1093/humrep/dey379
Advanced Access Publication on January 11, 2019

Reply: Ovarian response and its prediction are relevant

We thank our colleagues Dr Torrance, Dr Broekmans and Dr Mol for their interest in our study. In their letter, Torrance *et al.* question the value of predicting ovarian response in principle. This comes as a surprise, given that all three authors have themselves extensively been publishing on the prediction of high and low ovarian response (LOR) to ovarian stimulation (Hendriks *et al.*, 2005; Verhagen *et al.*, 2008; Broer *et al.*, 2013a, 2013b; Broekmans *et al.*, 2014; Hamdine *et al.*, 2015). While Torrance and Broekmans testify in their own work that 'accurate prediction of ovarian response prior to IVF is important' (Hamdine *et al.*, 2015) and that 'the predictability of ovarian response categories in antagonist co-treatment cycles is an important finding' (Broekmans *et al.*, 2014), they now have apparently come to the conclusion that ovarian response prediction is futile and irrelevant. This is even more surprising when considering that the same authors indeed utilize ovarian response prediction in order to design and conduct large interventional studies in both predicted low and predicted high responders (OPTIMIST trial: Oudshoorn *et al.*, 2017, van Tilborg *et al.*, 2017).

Regarding the clinical utility of predicting LOR we agree with Torrance *et al.* that at present no treatment has been shown to improve the prognosis for women with poor ovarian response. However, future interventional studies on new treatment options for LOR should use established prediction models for targeting subsets of patients by ovarian response. For daily clinical practice, prediction of LOR may furthermore be important in order to reduce treatment intensity and cost in those women, in which the therapeutic window for exogenous FSH stimulation has closed. Predicting a high ovarian response (HOR) to 150 μg corifollitropin alfa is undoubtedly of clinical relevance, since an FSH dose reduction can reduce burden, costs and risks without compromising outcomes (Nyboe Andersen *et al.*, 2017; Oudshoorn *et al.*, 2017). Regarding the cut-off, at which an individual should be considered at risk of HOR, we want to highlight that, firstly, only a small fraction of patients with HOR will indeed develop OHSS (Griesinger *et al.*, 2016), and secondly and more importantly, that our models allow the estimation of the HOR risk of an individual patient. This in turn allows the physician to make a case-by-case clinical judgment considering other patient specific risk factors. Finally, there is an obvious ex-ante interest of patients and doctors in the prospects of treatment by ovarian response (Sunkara *et al.*, 2011).

Regarding the methodological issues mentioned in their letter, we appreciate a fruitful discussion on the usage of well-suited methods. Some of their suggested methods definitely depict reasonable alternatives. Still, we consider some points as overly critical or unwarranted.

Torrance *et al.* raised the question why we 'chose single test accuracy measures instead of the more informative *c*-statistic and calibration plot'. We would like to point out that we explicitly presented area under the receiver operating characteristic curve (AUC) values—also called *C*-statistic or *C*-index—for all the compared models in Table II of our article.

They were puzzled why we 'still go on to develop new predictive models… [which] hardly outperform the Oehninger models'. While we openly discussed the non-superiority of the new models for the prediction of HOR, we did observe an improvement in the discriminative ability of the new models for the prediction of LOR. More importantly, novel models have to be developed in order to find out whether they can outperform the original ones, and the results of these scientific endeavors should be reported regardless of the outcome. On a side note, the main improvement of the LOR models was achieved by means of the log transformation of anti-muellerian hormone (AMH) levels, which is commented to be 'difficult to use in daily practice' by Torrance *et al.* We respectfully disagree on this point, computing a logarithm and an exponential function can be done with a simple hand-held calculator.

We wholeheartedly agree with Torrance *et al.* that the conventional validation route starts with an existing model, uses external validation and re-calibration if required, before entering clinical practice. We emphasize that in scientific practice, this is a continuing process that includes refinements and improvements of existing models. In that vein, our paper presents an external validation of the Oehninger models and at the same time suggests improvements of existing models that will, as we explicitly state, have to be externally validated in further studies.

They also raised the question why we 'chose to do a complete case analysis, which is known to introduce a risk of bias and reduce precision, instead of performing multiple imputation'. We agree that using multiple imputation of missing values might have been one way to go. However, complete case analysis might lead to biased results only if missing values are not missing completely at random, for which we did not see any indication. Certainly, complete case analysis comes at the

Downloaded from https://academic.oup.com/humrep/article-abstract/34/3/586/5288357 by guest on 06 March 2019

cost of reduced statistical power, but imputing missing values of explanatory variables by means of linear regression models, for example, increases the correlation between them, which is in fact not desirable. Particularly, missing values mainly concerned AMH and FSH of the same 34 patients, including four HOR cases and eight LOR cases. We considered these numbers of missing events and thereby minor loss in power as readily comprehensible. Instead we did not want to risk any imputation-induced effects, as for the HOR 1 model, for example, two important variables would have needed to be imputed for the same 34 patients.

Finally, they point out the risk to apply variable selection methods including nine potential predictive factors with only 32 events of LOR and 25 of HOR, respectively. We again agree on this point and strictly advise against using procedures like backward and forward selection blindly, especially with this small number of events. Although it is not mentioned in our article, we were aware of this potential risk and therefore ready to handle the findings of those methods with extreme care and to double-check each of the suggested predictor variables. In this case, there was actually no need for it as thoroughly described in our article. In fact, we even added important predictor variables to some models later on where we considered it appropriate, still yielding well-structured new models with fewer explanatory factors than those of Oehninger *et al.*, carefully following the 1 in 10 rule of thumb.

Conflicts of interest

M.O.S., T.L., I.R.K. and G.G. declare no conflicts of interest.

References

Broekmans FJ, Verweij PJ, Eijkemans MJ, Mannaerts BM, Witjes H. Prognostic models for high and low ovarian responses in controlled ovarian stimulation using a GnRH antagonist protocol. *Hum Reprod* 2014;**29**:1688–1697.

Broer SL, Dólleman M, van Disseldorp J, Broeze KA, Opmeer BC, Bossuyt PM, Eijkemans MJ, Mol BW, Broekmans FJ, IPD-EXPORT Study Group. Prediction of an excessive response in in vitro fertilization from patient characteristics and ovarian reserve tests and comparison in subgroups: an individual patient data meta-analysis. *Fertil Steril* 2013a;**100**:420–429.

Broer SL, van Disseldorp J, Broeze KA, Dolleman M, Opmeer BC, Bossuyt P, Eijkemans MJ, Mol BW, Broekmans FJ, IMPORT study group. Added value of ovarian reserve testing on patient characteristics in the prediction of ovarian response and ongoing pregnancy: an individual patient data approach. *Hum Reprod Update* 2013b;**19**:26–36.

Griesinger G, Verweij PJ, Gates D, Devroey P, Gordon K, Stegmann BJ, Tarlatzis BC. Prediction of ovarian hyperstimulation syndrome in patients treated with corifollitropin alfa or rFSH in a GnRH antagonist protocol. *PLoS One* **2016**;**11**:e0149615.

Hamdine O, Eijkemans MJ, Lentjes EW, Torrance HL, Macklon NS, Fauser BC, Broekmans FJ. Ovarian response prediction in GnRH antagonist treatment for IVF using anti-Müllerian hormone. *Hum Reprod* 2015;**30**:170–178.

Hendriks DJ, Mol BW, Bancsi LF, Te Velde ER, Broekmans FJ. Antral follicle count in the prediction of poor ovarian response and pregnancy after in vitro fertilization: a meta-analysis and comparison with basal follicle-stimulating hormone level. *Fertil Steril* 2005;**83**:291–301.

Nyboe Andersen A, Nelson SM, Fauser BC, García-Velasco JA, Klein BM, Arce JC, ESTHER-1 study group. Individualized versus conventional ovarian stimulation for in vitro fertilization: a multicenter, randomized, controlled, assessor-blinded, phase 3 noninferiority trial. *Fertil Steril* 2017;**107**:387–396.

Oudshoorn SC, van Tilborg TC, Eijkemans MJC, Oosterhuis GJE, Friederich J, van Hooff MHA, van Santbrink EJP, Brinkhuis EA, Smeenk JMJ, Kwee J *et al*, OPTIMIST study group. Individualized versus standard FSH dosing in women starting IVF/ICSI: an RCT. Part 2: The predicted hyper responder. *Hum Reprod* 2017;**32**:2506–2514.

Sunkara SK, Rittenberg V, Raine-Fenning N, Bhattacharya S, Zamora J, Coomarasamy A. Association between the number of eggs and live birth in IVF treatment: an analysis of 400 135 treatment cycles. *Hum Reprod* 2011;**26**:1768–1774.

van Tilborg TC, Torrance HL, Oudshoorn SC, Eijkemans MJC, Koks CAM, Verhoeve HR, Nap AW, Scheffer GJ, Manger AP, Schoot BC *et al*, OPTIMIST study group. Individualized versus standard FSH dosing in women starting IVF/ICSI: an RCT. Part 1: The predicted poor responder. *Hum Reprod* 2017;**32**:2496–2505.

Verhagen TE, Hendriks DJ, Bancsi LF, Mol BW, Broekmans FJ. The accuracy of multivariate models predicting ovarian reserve and pregnancy after in vitro fertilization: a meta-analysis. *Hum Reprod Update* 2008;**14**:95–100.

Markus O. Scheinhardt[1], Tamara Lerman[2], Inke R. König[1], and Georg Griesinger[2,*]

[1]*Institute of Medical Biometry and Statistics, University of Luebeck, University Hospital of Schleswig-Holstein, Campus Luebeck, Luebeck, Germany*

[2]*Department of Reproductive Medicine and Gynecological Endocrinology, University Hospital of Schleswig-Holstein, Campus Luebeck, Luebeck, Germany*

*Correspondence address. University Hospital of Schleswig-Holstein, Campus Luebeck, Ratzeburger Allee 160, 23538 Luebeck, Germany. Tel: +49-451-500-41951; Fax: +49-451-500-41954; E-mail: georg.griesinger@uni-luebeck.de

doi:10.1093/humrep/dey380

Advanced Access Publication on January 11, 2019

Downloaded from https://academic.oup.com/humrep/article-abstract/34/3/586/5288357 by guest on 06 March 2019