

A quantitative systems pharmacology workflow towards optimal design and biomarker stratification of atopic dermatitis clinical trials

Natacha Go (PhD)¹, Simon Arsène (PhD)¹, Igor Faddeenkov (MSc)¹, Théo Galland (MSc)¹, Shiny Martis B. (PhD)¹, Diane Lefaudeux (MSc)¹, Yishu Wang (PhD)¹, Loic Etheve (PhD)¹, Evgueni Jacob (MSc)¹, Claudio Monteiro (PhD)¹, Jim Bosley (PhD)¹, Caterina Sansone (PharmD)², Christian Pasquali (PhD)², Lorenz Lehr (PhD)², and Alexander Kulesza (PhD)¹

¹Novartis, Lyon, France

²OM Pharma, Meyrin, Switzerland

Funding: Novartis and OM Pharma funded this study.

Abstract:

Background. The development of atopic dermatitis (AD) drugs is confronted by many disease phenotypes and trial design options, which are hard to explore experimentally.

Objective. Optimize AD trial design using simulations.

Methods. We constructed a quantitative systems pharmacology (QSP) model of AD and standard of care (SoC) treatments and generated a phenotypically diverse virtual population whose parameter distribution is a) derived from known relationships between AD biomarkers and disease severity and b) calibrated using disease severity evolution under SoC regimens.

Results. We applied this workflow to the immunomodulator OM-85, currently being investigated for its potential use in AD, and calibrated investigational treatment model with the efficacy profile of an existing trial (thereby enriching it with plausible marker levels and dynamics). We assessed the sensitivity of trial outcomes to trial protocol and found that for this particular example, a) the choice of endpoint is more important than the choice of dosing-regimen and b) patient selection by model-based responder enrichment could increase the expected effect size. A global sensitivity analysis reveals that only a limited subset of baseline biomarkers is needed to predict the drug response of the full virtual population

Conclusion. This AD QSP workflow built around knowledge of marker-severity relationships as well as SoC efficacy can be tailored to specific development cases so as to optimize several trial protocol parameters and biomarker-stratification and therefore holds promise to become a powerful model-informed drug development tool.

Key Messages:

- Disease and treatment models can quantify pre-existing knowledge about complex immune diseases such as atopic dermatitis and drug's efficacy data under one common umbrella.
- Embedding QSP models into trial simulation setup can give insight into clinical trial optimization.
- Complex QSP models can help with patient selection and biomarker identification.

Capsule Summary: This study shows the relevance of QSP model and computer simulations in assisting clinical development in the field of atopic dermatitis by assessing the impact of trial protocol on treatment effect and guiding biomarker programs.

Keywords: Atopic dermatitis, Trial design, Trial optimization, Mathematical modeling, Biomarkers, Best responder, In silico approaches, Immunomodulation, Bacterial lysates

Abbreviations:

AD Atopic dermatitis

SoC Standard of care

TCS Topical corticosteroids

ODE Ordinary differential equation

QSP Quantitative systems pharmacology

PBPK Physiologically based pharmacokinetic

PD Pharmacodynamics

AB Absolute benefit

SCORAD Scoring atopic dermatitis

EASI Eczema area and severity index

1 Introduction

Atopic dermatitis (AD) is a chronic inflammatory condition of the skin characterized by recurrent eczematous lesions and intense itch which can profoundly impair quality of life¹⁻⁶. The pathophysiology of AD is complex and involves local and systemic immune dysregulation, genetic susceptibility, environmental factors and microbiome effects^{7;8}. AD is not a homogeneous disease encompasses a variety of endotypes and phenotypes in the different age, ethnic, etc. groups (see the work cited in Facheris et al.⁹). The road to novel therapeutics¹⁰ is tortuous, due a combination of two factors:

- a) a partial understanding of the molecular and cellular mechanisms driving disease severity and treatment effect in individual patients^{11;12}, making it difficult to translate preclinical results into the clinic and even more so to estimate the quantitative effect in a given population.
- b) the complex clinical management of AD^{13;14} which opens many up many potential choices in trial design such as rules for rescue treatment or definition of induction or washout periods¹⁵.

Despite advancement in our understanding of epidemiology, biomarkers, endotypes, prevention, and comorbidities¹⁶ as well as *in vitro*, *in vivo* and *in silico* approaches developed for the investigation of human AD pathogenesis¹⁷, the precise relationships between intrinsic immunological features, extrinsic factors, disease severity and heterogeneous treatment effect remains to be established. *In silico* approaches based on mechanistic models can be powerful tools to explore trial design options^{18;19} in a context, such as AD, where high heterogeneity makes that, in addition to questions on protocol and dosing-regimen, the strategies involving patient selection, involving for example local or systemic biomarkers, are of utmost importance²⁰. There are several existing mathematical models for AD¹⁷, notably highlighting the importance of temporality in skin barrier function, immune responses, and impact of environmental stressors²¹ as well as the identification of effective biologic drugs combination for single-drug non-responders²². These studies have demonstrated that a systems modeling approach can help resolving mechanistic questions in AD drug development. Still, the question of how to quantitatively translate these insights into model-informed clinical

trial design remains open. Within model-informed drug development, QSP models are being increasingly considered, both for internal and regulatory decision making²³. QSP - including within a trial simulation paradigm - has produced numerous published examples and regulatory grade evidence across disease areas where protocols and in particular the dosing-regimen aspects have been informed^{23;24}, but the use of QSP for informing predictive biomarker programs have been mostly limited to generating plausible dynamics of important biomarkers in different tissues and did not yield quantitative and actionable information on clinical applications such as e.g. a biomarker identification trial. This gap motivated us to combine a QSP modeling approach for an investigational treatment in AD with a trial simulation strategy aiming at both, quantifying the relationship between trial design options and the expected efficacy profiles while taking into account the patients' immune profiles and thereby addressing biomarker related trial design choices.

In this work, we consider the investigational treatment OM-85 as a use-case. The oral immunomodulator OM-85 is a bacterial lysate which modulates the immune system notably by restoring the Th1/Th2 balance²⁵ and is thus being investigated for various atopic conditions²⁶. OM-85 is approved in several countries for the prevention of respiratory tract infections²⁷ and has shown some promising results in pediatric AD as add-on treatment²⁸. It is now being tested for early moderate AD in an ongoing trial (NCT05222516) where the immune heterogeneity of the included population can represent a challenge. We developed a novel mechanistic model of skin immune dysregulation in AD including the mechanisms of action of emollients and topical corticosteroids (TCS), combined with a previously developed model of the administration, pharmacokinetics, and mechanism of action of OM-85 (originally applied in combination with a multiscale respiratory tract infection disease model²⁹). We calibrated this QSP model with quantitative relationships in skin biomarkers^{30–35}, clinical data for standard of care (SoC)^{22;36–39} and OM-85 treatment efficacy²⁸. As a plausible application example, we used this QSP model to run trial simulations for virtual AD populations in order to obtain both patient- and trial protocol-specific predictions of drug efficacy profiles. We found that trial efficacy is expected to be insensitive to the exact dosing regimen but can display sensitivity with respect to choice of severity endpoint (relative vs absolute), which thus

104 appears to be an option for further trial optimization. We then show how a dimensionality reduction
105 approach can be used to inform a biomarker-based stratification strategy: global sensitivity analysis
106 identifies a reduced number of baseline biomarkers in the virtual population, that reproduce the
107 reference data well, and thus can be used to select patients based on their predicted treatment effect.
108 We finally show that in the chosen use case, both protocol design and patient selection with results
109 of the model do show optimization potential and therefore may ultimately support decisions about
110 clinical (predictive) biomarker programs.

2 Methods

2.1 QSP workflow: combining systems immunology and trial simulations

Our QSP workflow (modeling and simulation approach) is based on a knowledge and mechanism-driven mathematical model focusing on the immunology of AD, built on the back of a system of ordinary differential equations (ODEs) and used in a setup to simulate clinical trials (also referred to as a quantitative disease-drug-trial model⁴⁰ or *in silico* clinical trials). A virtual population approach captures between-patient variability, where model parameters are described by statistical (co)distributions rather than scalar values.

2.2 QSP model description

The QSP model (Fig. 1) can be broken down into three components: a) the disease model describing the skin barrier integrity and the immune system and linking baseline biomarkers levels to AD severity scores, b) the SoC treatment models accounting for the administration and mechanisms of action of emollients and TCS, and c) the investigational treatment model describing the administration, pharmacokinetics, and mechanism of action of OM-85. The disease model focuses on selected quantitative aspects of skin immune dysregulation in patients with AD while not taking into account pathogenesis or natural evolution of AD. This allowed us to consider that, without treatment, the disease state of a given patient with a given AD severity can be characterized by stationary dynamics of the biological entities. Under treatment however, the stationary immune dynamics of the disease model (*i.e.* at baseline) are perturbed and the AD severity evolves. The calibration of the QSP model was performed following the core approach for complex disease-treatment-trial models with heterogeneous data as described in Palgen et al.⁴¹.

2.2.1 Atopic dermatitis disease model

The disease model covers the main mechanisms at the cellular scale involved in the modulation of the skin barrier integrity and immune dysregulation in response to environmental stressors. In particular,

it captures the mechanisms driving the dynamics of the major biomarkers of AD lesional skin including filaggrin, *S. aureus* in the skin microbiome, and major soluble immune biomarkers (mainly cytokines and chemokines characterizing the type-1, 2, 17, 22, and regulatory responses). While for typical pharmacometrics analyses, the principle of parsimony guides the selection of the model structure⁴², our mechanistic model is complex and combines a heterogenous set of information from different sources. Its architecture consists of a system of ODEs representing the dynamics of 21 variables and involves 139 parameters (Fig. 1, Supplementary Information S1.1).

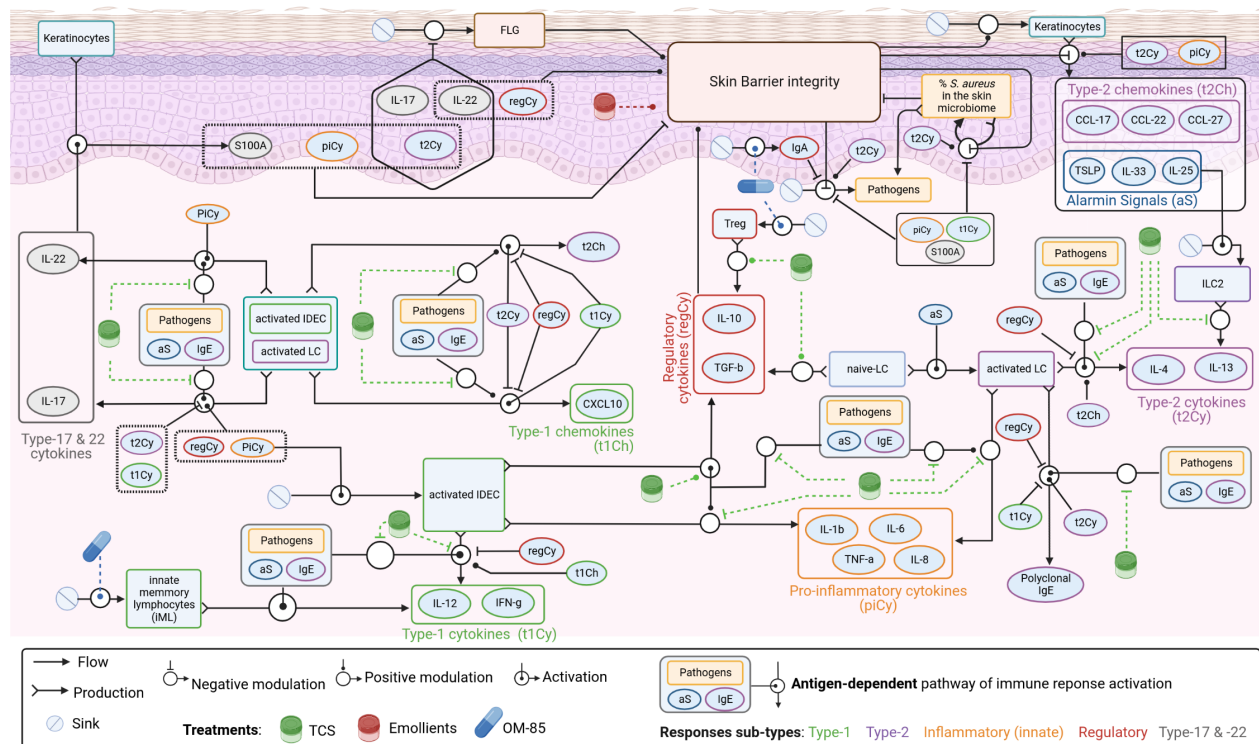


Figure 1. QSP model schematic. Exhaustive representation of the interplay between the skin barrier and the skin immune system as formalized in the AD disease model, mechanisms of actions of SoC treatments (TCS and emollients) and the investigational treatment OM-85.

The main mechanisms included in the disease model can be summarized as follows. The skin barrier integrity is modulated by filaggrin levels⁴³, harmful *S. aureus* colonization^{6;44}, and skin immune dysregulation^{6;44–49}. In return, conjointly with immune dysregulation, skin barrier integrity modulates the level of pathogen infiltration^{50–52} and determines the production of alarmins (TSLP, IL-25, IL-33). Alarmins activate the innate immune response, notably Langerhans cells (LC) and

inflammatory epidermal dendritic cells (IDEC),^{47;52–54} both in antigen-independent and -dependent pathways (via alarmins or specific IgE, respectively)^{55;56}. Both LC and IDEC synthesize innate pro-inflammatory factors (IL-1 β , IL-6, TNF- α , IL-8)^{53;56} as well as type-17/22 immunity proteins (IL-17, IL-22, and S100A)^{43–45;48;57}. Activated LCs – in context of AD⁵³ – induce the polarization of the adaptive response towards type-2 (represented by IL-4, IL-13 and IgE) through the production of type-2 chemokines (CCL-17, CCL-27 and CCL-22), while IDEC polarize it towards type-1 (represented by IL-12 and IFN- γ) through the production of type-1 chemokines (CXCL-10)⁵⁶. The interconnected feedback loops of the type-1 vs type-2 response converge to an overall type-2 skewing in AD, which in return is mixed with growing type-1 response with increasing AD severity^{53;56;57}. In AD context, regulatory cytokines (IL-10 and TGF- β) are produced by regulatory T cells (Treg) and dendritic cells (naive LC and activated IDEC with high level of Fc ϵ RI on their cell surface)^{47;53;56;58} and limit the innate inflammation as well as type-1 and type-2 responses^{44;58;59}. AD severity (evaluated with SCORAD⁶⁰ and EASI⁶¹) is phenomenologically linked to the level of skin barrier integrity and pathogen infiltration²².

We make the fundamental assumption that, without treatment, the disease model represents the skin immune system at equilibrium (Supplementary Information S1.2.1). By computing and applying the mathematical conditions for equilibrium, we are able to derive a large number of parameters (reducing the set of free parameters from 113 to 23, Supplementary Information S1.2.1, Supplementary Information Table S7). We can therefore inform the steady-state values of the thirteen biomarkers included in the model as a function of AD severity for individual patients based on a collection of datasets from the literature (Supplementary Information S1.2.2, Supplementary Information Table 4). Note that the entire dynamic system is still underinformed meaning that there are still parameters for which several values satisfy the conditions for equilibrium. For this reason it is important to also constrain the dynamical behavior upon perturbation of the equilibrium, i.e. under treatment (where target engagement and pathway modulation will drive this system out of equilibrium) as described in what follows.

2.2.2 Standard of care treatment model

The SoC treatment comprises a combination of emollients and different TCS (Fig. 1). The application of emollient ameliorates the skin barrier integrity^{3;62;63} while TCS inhibit type-1, type-2 and pro-inflammatory responses and amplify the regulatory response by modulating respective cytokines^{64–68}. We disregarded topical drug pharmacokinetics by considering a standardized bioavailable concentration throughout a day of treatment regardless of duration of application, dosage and intra-daily frequency. We also neglected the potential systemic effects given the very low systemic absorption of modern TCS^{69;70}. We calibrated the SoC treatments against an aggregated dataset of 5 studies accounting for various combinations and potencies (Supplementary Information S1.2.3, Supplementary Information Table S6). The set of calibrated parameters include those driving SoC mechanisms of action but also those controlling AD pathophysiology left undetermined after applying equilibrium constraints (Supplementary Information Table S7). Finally, we used the placebo arm (only emollients + TCS) reported in our reference dataset for the investigational treatment (Bodemer et al.²⁸) to assess the robustness of model prediction for SoC treatments.

2.2.3 Investigational treatment (OM-85) model

For OM-85 administration and effect, we re-used the model reported in Arsène et al.²⁹ where, as described in the context of respiratory tract infections prophylaxis, the oral administration of OM-85 triggers the activation and proliferation in the intestinal Peyer's Patches of reprogrammed type-1 innate memory like cells^{71–73}, regulatory T-cells^{74–76} and polyclonal IgA producing plasmablasts^{77;78}, which further disseminate via the systemic circulation into the inflamed lung tissues, due to the presence of chemokines and homing receptors. In the context of AD, the fundamental hypothesis is that that similar cells infiltrate the inflamed skin tissues (the gut-skin axis as a central mechanistic hypothesis), which is backed up notably by many reports of co-expression of gut, skin and lung tissue homing markers^{52;58;79;80} and is consistent with the reported effect of OM-85 in AD²⁸. To take this hypothesis into account, the model reported in Arsène et al.²⁹ is supplemented with a layer describing skin-homing of activated immune effectors (Fig. 1, Supplementary Information Figure

S1).

2.3 Analyses

2.3.1 Treatment effect

The treatment effect can be quantified *in silico* by the individual absolute benefit (AB, see Boissel et al.⁸¹) which is defined as the difference of SCORAD at 6 months for a specific virtual patient with placebo vs when treated (in our idealized clinical setting each virtual patient can be its own control).

2.3.2 Average treatment effect

The average treatment effect is the group equivalent of treatment effect, defined as the difference of the mean severity score at 6 months in the placebo vs treatment arm ($\overline{AB}$, see e.g. Bodemer et al.²⁸). We obtain confidence intervals around predicted values by using a bootstrapped approach (using 100 samples, each with 85 virtual patients for the placebo and for the treatment arms). For the sensitivity assessment of the trial results with respect to the choice of main endpoint, we used four alternative definitions for the average treatment effect, either based on SCORAD, EASI, relative SCORAD or relative EASI. For relative SCORAD or EASI, the difference in severity scores at 6 months is normalized by the mean severity score at 6 months in the placebo arm.

2.3.3 Power and sample size

We compute empirical statistical power through bootstrapping⁸² using Students t-test (alpha risk = 0.05) for various sample sizes. This allows us to estimate the required sample size to reach a statistical power of 0.80(Supplementary Information S.1.4).

2.3.4 Recruitment effort

We define recruitment effort as the total number of patients to be screened given selection criterion: we compute it by dividing the estimated sample size for the corresponding selection criteria by the

222 percentage of virtual patients matching the selection criteria in the reference virtual population, for
223 which the distribution of patient characteristics is assumed to be realistic.

3 Results

3.1 Reproduction of a placebo-controlled clinical trial

In order to quantitatively align our QSP model with a specific clinical development case, we adapted a simulation and analysis protocol to represent an existing trial (Bodemer et al.²⁸ [NCT05222516]) and used the reported efficacy profiles to calibrate the remaining undetermined parameters (controlling the effect of OM-85). This simulation protocol corresponds to a placebo-controlled parallel two-arm AD trial for a virtual population of pediatric subjects treated with OM-85 for 9 months and assessing AD severity (SCORAD) during the study (Table 1). In order to match the longitudinal efficacy data for OM-85, we calibrated the three parameters controlling the effects of immune effectors activated by OM-85 against the average evolution of disease severity in the treatment arm (all the other model parameters being already calibrated: see SoC treatment and Arsène et al.²⁹). As illustrated in Fig. 2, average value and variability of disease severity evolution in absolute and in relative values, and value at 6 months (Fig. 2 A-B, C-D, E-F respectively) are well reproduced, for both placebo and treatment arms.

Virtual population	Age	$\ln(\text{Age}) \sim \mathcal{N}(0.3, 1.8) \text{ \& } 0.6 \leq \text{Age} \leq 8$			
	Gender	$\sim \mathcal{B}(0.5)$			
	SCORAD	$\ln(\text{SCORAD}) \sim \mathcal{N}(3.7, 0.27) \text{ \& } 15 \leq \text{SCORAD} \leq 95$			
	Biomakers	function*(SCORAD) + $\mathcal{N}(\text{mean}^*, \text{sd}^*)$			
Administration regimen		Week 0 – 1	Week 1 – 2	Week 2 – 3	Week 4 – 36
	Emollients	Every 1 day			
	TCS	Every 1 day	Every 2 days	Every 3 days	Every 3 days (*)
	OM-85	Every 1 day			
Arms	Placebo	Emollients + TCS			
	Treatment	Emollients + TCS + OM-85			
Measurements		Baseline (day 0)		Follow-up: month 1, 3, 6, 9	
	Data ²⁸	SCORAD, flares			
	Model	SCORAD, EASI, biomarkers		SCORAD, EASI	

Table 1. Simulation protocol for in silico clinical trials mimicking the randomized reference clinical trial in Bodemer et al.²⁸ [NCT05222516]. Virtual population (2000 patients / arm), administration regimen, arms, and measurements. ★ Calibrated.

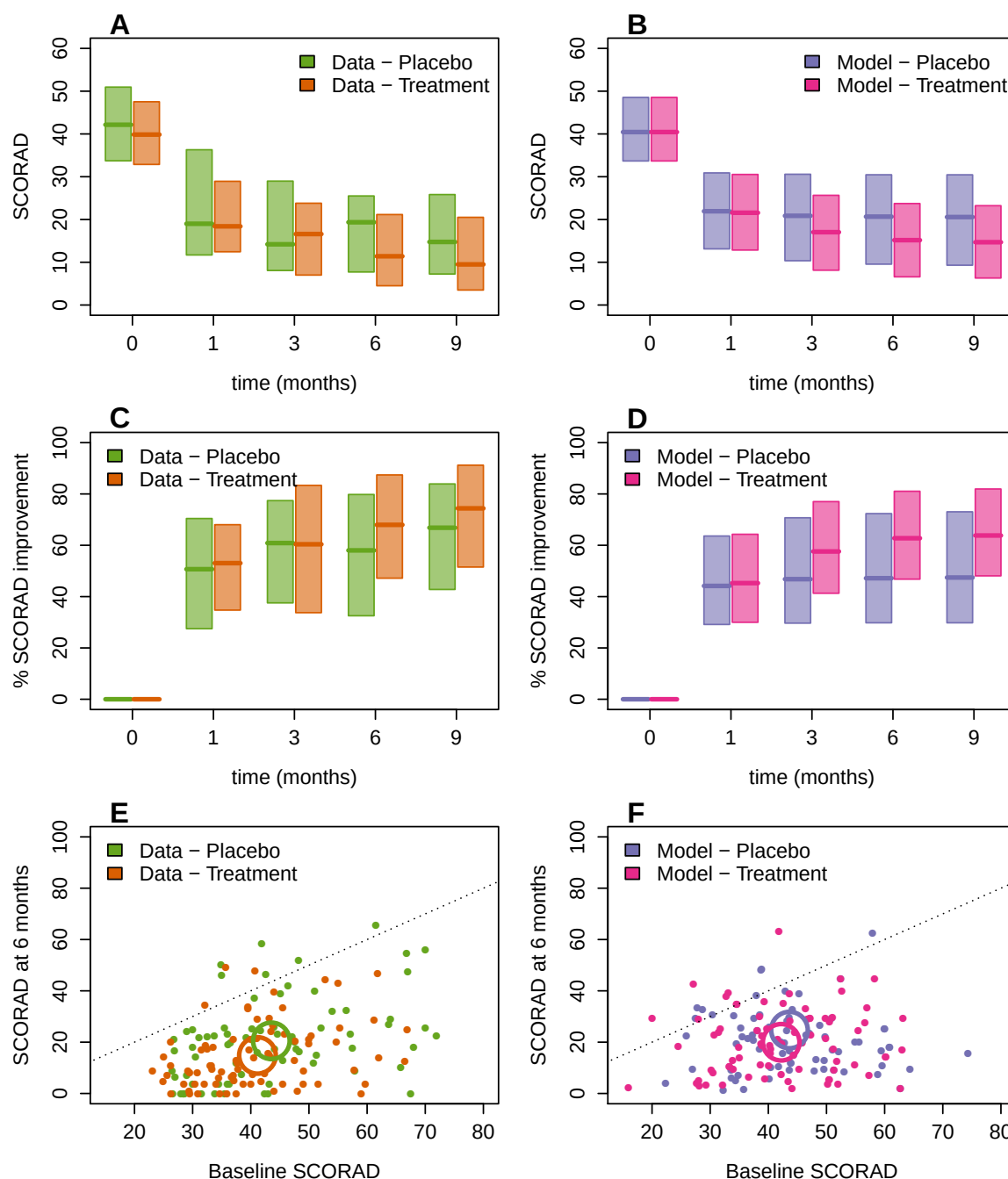


Figure 2. Reproduction of a randomized placebo-controlled clinical trial. A – B: Evolution of AD severity (SCORAD, median: line, quartiles: box) of treatment vs placebo arms from the real trial (Bodemer et al. ²⁸) (A) and virtual population (B) at observation time points. C – D. Evolution of percentage improvement of AD severity (SCORAD, median: line, quartiles: box) of treatment vs placebo arms from the real trial (Bodemer et al. ²⁸) (C) and virtual population (D) at observation time points. E – F: Individual (points) and average (circles) treatment effect represented by the SCORAD at 6 months compared with baseline SCORAD from the real trial (Bodemer et al. ²⁸) (E) and randomly sampled virtual patients matching the baseline SCORAD for placebo and treatment arms (F).

In particular, SCORAD at 6 months for data vs model are Q1: 7.7 vs 10.2 median: 19.4 vs 21 Q3: 25.5 vs 31. The average calibration error in median SCORAD for the placebo arm is 4.1 (range: 1.7 7.1), and 3.2 (range: 0.6 5.8) for the treatment arm. These results show that a predicted impact of protocol changes to SCORAD higher than 5 points can be considered to exceed the efficacy prediction error of our QSP model.

3.2 Assessment of trial results sensitivity to the trial protocol

Trial protocols are complex and contain a high number of degrees of freedom. As such, their impact on efficacy are typically difficult to assess empirically. To address this difficulty, we intend to show that *in silico* trials based on our QSP model can allow clinicians to systematically test variations in trial protocol and conclude on possible trial outcomes. We therefore explored potential design alternatives for the trial reproduced in Section 3.1 (see also Table 1) in order to identify the trial protocol parameters to which the average treatment effect is most sensitive. We report the average treatment effect and sample size estimated for given power in Fig. 3. Overall, the average treatment effect shows a clear sensitivity to variations in dosing regimen of OM-85 administration (Fig. 3A-D) or in the choice of the endpoint (Fig. 3E-H) whereas the impact of SoC treatments parameters (duration of the TCS induction phase, TCS potency, and TCS administration frequency) appear to be negligible (Supplementary Information S2.1).

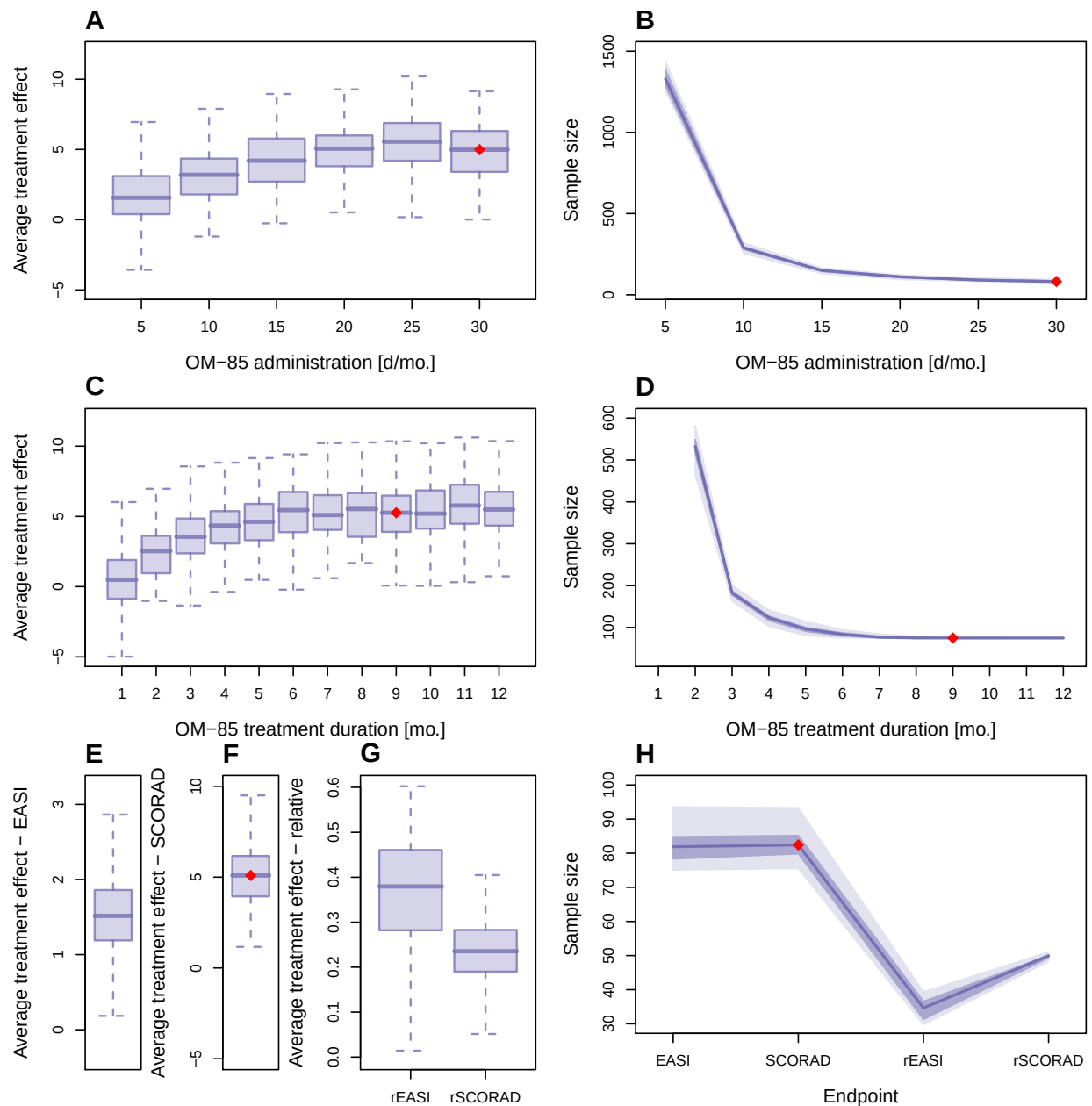


Figure 3. Sensitivity of trial results to trial protocol. Impact of (A – B) OM-85 administration frequency, from 5 to 30 days / month with a 5-day incremental step; (C – D) OM-85 treatment duration, from 1 to 12 months with a 1-month incremental step; (E – H) Endpoint: EASI, SCORAD, relative (*i.e.* normalized by the baseline severity) EASI (rEASI) and SCORAD (rSCORAD). The red dot represents the clinical trial settings of the study of Bodemer et al. ²⁸.

The average treatment effect increases with the frequency of OM-85 administration with the best performance (Fig. 3 A: average treatment effect: 5, B. sample size: 83) being reached with a daily OM-85 administration as in Bodemer et al. ²⁸. Compared with 5 days of OM-85 administration per

month, the average treatment effect is 2 fold higher and the sample size 17 fold smaller. The efficacy also increases with the treatment duration until 6 months of treatment before reaching a plateau (Fig. 3 C: plateau average treatment effect 5.4, D: plateau average sample size 77). Compared with a single month of treatment, the average treatment effect is 10 fold higher and the sample size 6 fold smaller. Finally, our results indicate that the best endpoint is the relative EASI (Fig. 3E-G) for which the power (Fig. 3H) and sample size (Fig. 3I) are minimal: 34 patients are required compared with 83 when using absolute SCORAD. Note that average treatment effect appears to be more sensitive to the choice of an absolute vs relative endpoint than to the choice of scoring system (SCORAD or EASI, Fig. 3E).

3.3 QSP model-assisted early biomarker identification

In this section, we showcase how our QSP model can support a biomarker program early in clinical development with the aim of tailoring the target population by inclusion/exclusion criteria. This strategy is relatively common in AD: for example, focusing on higher disease severity strata can reduce population heterogeneity and potentially increase measured clinical benefit (explained e.g. by a larger need for treatment) but at the expense of targeting a more narrow subpopulation of AD patients¹. Selection by phenotype and/or biomarkers may also help with treatment effect stratification and development of companion diagnostics already had success in other disease areas, e.g. in targeted cancer therapy⁸³. However in the context of AD, the identification of predictive biomarkers seems to be difficult and to date, stratification by phenotypes (or associated markers) was not yet successful in stratifying treatment effect⁸⁴ and only a handful of treatment-specific molecular biomarkers have been suggested so far with targeted therapies (see⁸⁵). In part, this is due to the fact that identification of biomarkers predictive of the treatment effect requires the existence of rich datasets which include treatment effect and a panel of marker candidates. In the absence of such data, QSP models, acknowledged by regulatory bodies⁸⁶ (especially in pediatric populations), have been suggested as promising tools to guide biomarker identification because they can leverage knowledge about the involved pathophysiological mechanisms thereby encoding information on

how disease pathways respond to biomarker change.

3.3.1 Enriched synthetic trial data generation

We make use of the following two capabilities of the disease model: a) the uncorrelated random variations in the 13 immune biomarker dimensions ([Supplementary Information Table S4](#)) in the virtual population which are in line with known marker variability and correspond to a broad diversity of AD phenotypes b) the alignment of the in silico clinical trial with historical trial data (Section 3.1). This allowed us to predict hypothetical, but biologically plausible, individual treatment effect for each virtual patient after OM-85 treatment including detailed immune cellular and soluble marker dynamics. This procedure augmented the existing demographics and disease severity data into a synthetic but highly immunologically enriched individual patient data set (Fig. 2). In the following sections, we then used this synthetic data to investigate various possible biomarker strategies for treatment effect stratification. Theoretically, there exists a range of options: from the use of disease severity strata to the use of thresholds on baseline markers levels (e.g. informed by a classifier based on linear or logistic regression and receiver operating curve validation) or of entire panels of quantitative marker levels (e.g. using treatment effect predicted from a model). This last strategy is known from other disease areas: for example, many metabolic disease trials employed stratification by insuline resistance as defined by HOMA-IR which uses fasting insulin and glucose levels and which was originally formulated as a rather complex non-linear ODE model⁸⁷.

3.3.2 Disease severity shows only limited potential for treatment effect stratification

In line with the common practice of selecting AD trial populations by severity, we investigated if baseline severity was a good predictor for the OM-85 effect in the reproduced and enriched reference trial data. As a first simple test, we performed a linear regression of the OM-85 effect with respect to virtual patient's baseline disease severity, which resulted in a poor fit ($R^2 = 0.29$). In order to exclude that the too simplistic picture of a linear regression confounds a severity stratification effect, we also employed our QSP model, but using a simplified virtual population

where only individual baseline disease severity is varied ([Supplementary Information S2.2](#)). To better explain this approach, an analogy to mixed effect modeling⁸⁸ is useful: in the reference virtual population, severity would be a random effect while all biomarkers would be mixed (with a severity-dependent mean and a random component). In the simplified severity-only population, all biomarkers' fixed effects are kept but the random components are discarded. This approach performs poorly in comparison to the treatment effect values from our synthetic reference data ($R^2 = 0.35$, [Supplementary Information Figure S8.A](#)), similarly to the linear regression. In line with that result, selecting a more severe patient subpopulation in our synthetic reference data is associated with slightly increased mean average treatment effect ([Supplementary Information Figure S8.C1](#)) as well as a predicted smaller sample size ([Supplementary Information Figure S8.C2](#)), but the restriction of the eligible patient fraction grows exponentially with the selection criteria ([Supplementary Information Figure S8.C3](#)). Selecting patients by baseline disease severity is therefore not the most promising avenue to responder enrichment in this example. This interpretation, however, has to be considered with caution since even though restricting the target population might be acceptable in some cases, clinical management aspects may become limiting for patient selection and more severe AD may have distinct pathophysiological mechanisms.

3.3.3 Workflow for using our QSP model for low-dimensionality biomarker strategy

Consequently, in view of the unsatisfactory results of stratification by disease severity we therefore proceeded to take into account the coupled and non-linear nature of biomarker dynamics implemented into our QSP model. While in principle one could use the QSP model to predict treatment effect (and other output variables) using all biomarkers as inputs, this strategy may not be realistic in practice. Indeed, the need of comprehensive data to provide a value for all inputs is in contrast with the ethical burden of taking patient samples as well as economical considerations (e.g. extensive proteomic analyses⁸⁹). Furthermore, the decision of committing to such a large data acquisition plan is difficult to take without some confidence in the fact that these data could efficiently stratify the patient population. Specific parts of our treatment-agnostic model may be less important for the

effect of a specific treatment (such as OM-85, which acts via increase of the regulatory response, notably). Therefore, we sought a more parsimonious approach.

A preliminary investigation using typical model reduction technique fitting a series of generalized linear models ($R^2 = 0.2$, [Supplementary Information S2.3](#) and [Supplementary Information Figure S9.A](#)) did not show promising trade-off between simplification potential and accuracy when simplifying the QSP model structure. We therefore aspired to develop a workflow where not the QSP model, but the overall in silico trial approach is simplified. For this, we reduced the virtual population by replacing some biomarker values by reference values (which depends on disease severity) with the idea of not drastically altering treatment effect, the primary outcome. Again, with a mixed effect modeling paradigm in mind, this amounts to removing as many random effects as possible without degrading the model's goodness of fit. Such dimensional reduction approach has several advantages: a) it increases the chances for validation with experimental data and b) it enables the integration with a biomarker identification program in a realistic clinical setting with a limited number of biomarkers per test. This strategy is similar to feature selection techniques that have been suggested earlier to interface machine learning with QSP⁹⁰.

Using the example of OM-85 versus placebo as add-on therapy on top of TCS and emollients, we therefore analyzed the contribution of the set of inputs in treatment effect variability. The result of this is the ranked list of markers that need to be supplied to the model (either from clinical data or the virtual population as a plausible theoretical counterpart of such data) in order to predict treatment effect. In practice, we performed an ANOVA-based global sensitivity analysis^{91;92} of treatment effect to baseline biomarkers using a virtual population of 2000 patients characterized by a unique baseline SCORAD ([Supplementary Information Figure S11](#)), which allowed to quantify the main effect of each biomarker and the effects of up to between four biomarker interactions. Ordering the biomarkers by their total effect (biomarker alone and sum of the interactions with any other biomarker) shows that each of the top four - alarmins, IL-22 cytokines, regulatory cytokines, and type-2 cytokines - explains more than 20% of the variance (and 83% in total) (Fig. 4B). Based on their main effect only (and not higher order interactions), they also exhibit the highest influence

(8 to 15% each, 46% in total [Supplementary Information Figure S12.A](#)). We thus selected these lead biomarkers to predict treatment effect in the reduced input dimensionality approach. We then evaluated how this input dimensionality reduction approach performs, by confronting treatment effect values predicted with the reduced-variability vs the reference virtual population (Fig. 4A). While this approach leads to slight overestimation of treatment effect (Fig. 4C), the accuracy of the predictions is satisfactory (Fig. 4D, RMSE = 2.6, $R^2 = 0.63$). Indeed, the prediction error (limited vs. full approach) is lower than 5 (SCORAD) for 94% of the virtual patients and the average error is 1.1 (min: 0.0002, Q1: 0.45, Q3: 2.23, max: 24.5). This prediction error is notably lower than measurement errors typically reported (normally distributed with a SD of 5⁹³).

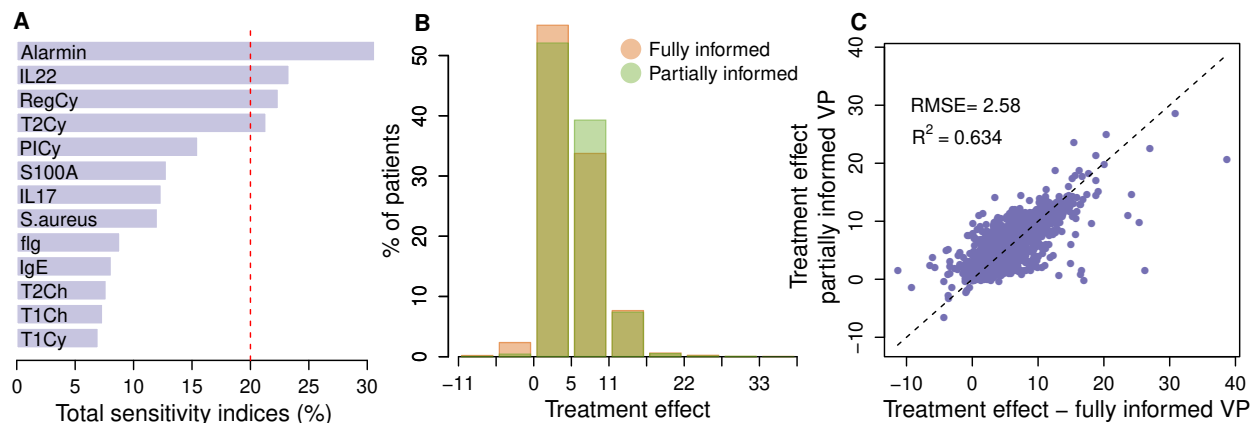


Figure 4. Low-dimensionality biomarker strategy. A. Global sensitivity analysis quantifying the (total: main + interactions) influence of baseline biomarkers on the treatment effect variance; selected key biomarkers have a sensitivity index above the red dashed line (20 %). B. Treatment effect distribution from the fully vs partially informed virtual population. C. Performance of the dimensionality reduction approach: comparison of treatment effect predicted by the fully vs partially informed virtual population.

We finally assessed the potential of using the biomarkers identified with the input dimensionality reduction approach for trial protocol optimization. In particular, we tested the effect of selecting virtual patients with required increasingly high treatment effect (predicted upon baseline marker selection); and we constructed subpopulations (including all patients – without selection – to the top 5% responders in incremental steps of 5%). For assessing this strategy's performance, we report three indicators in Fig. 5: average treatment effect, sample size and statistical power. Selecting the top 50% responders results in 50% improvement of the average treatment effect (Fig. 5A), 50%

378 reduction of the average sample size (Fig. 5B), and 31% improvement of the average chance of
 379 success of the trial (Fig. 5C). The maximal improvement for the average treatment effect (with
 380 respect to reference trial, enrolling only the top 5% predicted responders) is roughly 150% (Fig. 5A),
 381 the maximal reduction of the average sample size is 75% (Fig. 5B), and the maximal improvement of
 382 the average chance of success 33% (Fig. 5C). We also evaluated the impact of patient selection on the
 383 recruitment effort, defined as the total number of patients to be screened (Fig. 5D). The recruitment
 384 effort can be seen as a metric of the optimal selection criteria, as in the clinics, part of the cost would
 385 scale with the number of patients to screen. Our results suggest the optimal selection criterion lies
 386 between 25% and 60% of best responders, where the recruitment effort is the lowest, with a 19%
 387 average decrease compared to no selection. The recruitment effort exponentially increases with
 388 the selection stringency when selecting the top patients up to the top 60%, representing up to 300
 389 patients to be screened for the top 5% selection criterion, *i.e.* 275% more than without selection.

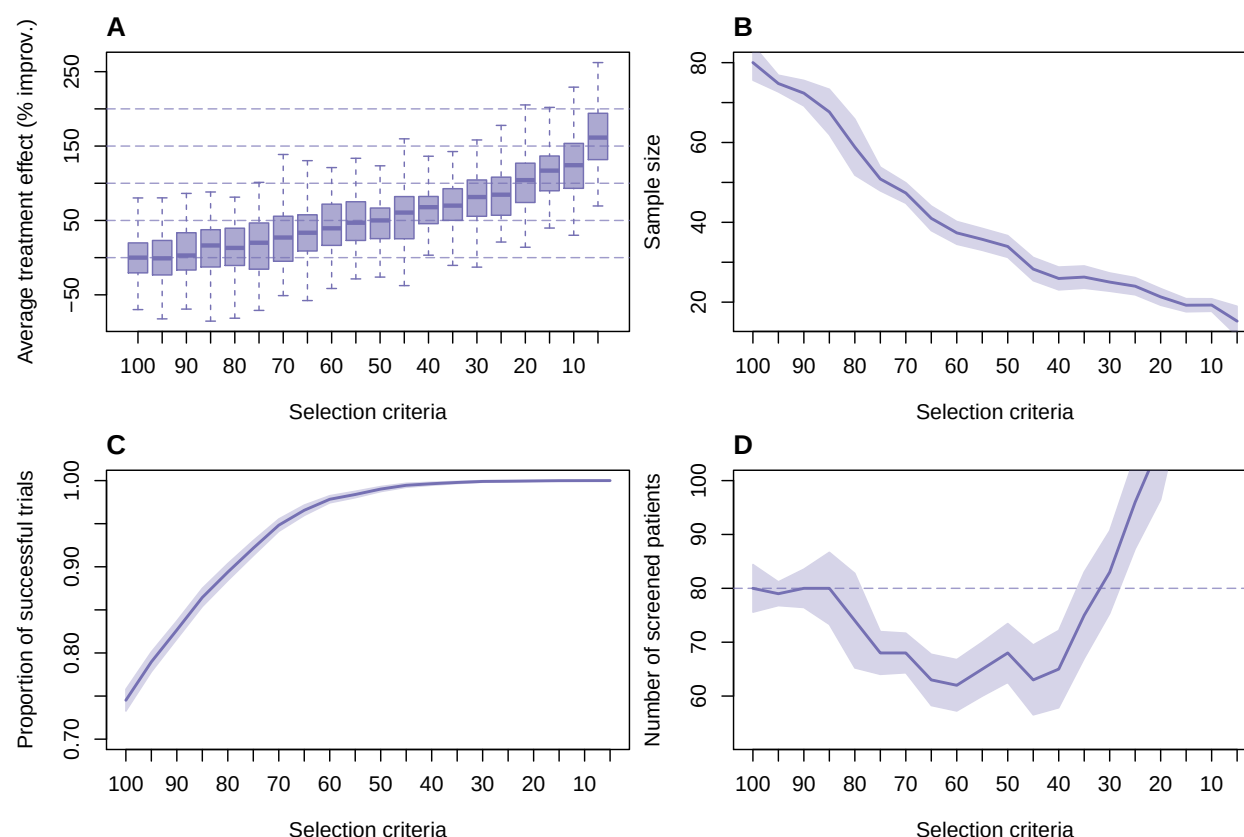


Figure 5. Clinical trial optimization based on predicted responders selection. A. Average treatment effect improvement (in %) vs increasingly strict selection criteria (top % responders): median (line), interquartile range (box), whisker (dashed line). B. Sample size distribution (obtained by bootstrapping over 100 samples) vs increasingly strict selection criteria: mean (line) and standard deviation (colored area). C. Statistical power (*i.e.* proportion of successful trials) distribution (obtained by bootstrapping over 100 samples) vs increasingly strict selection criteria : mean (line) and standard deviation (colored area). D. Recruitment effort (total number of patients to be screened during the recruitment process) distribution) vs increasingly strict selection criteria: mean (line) and standard deviation (colored area). Note that the y-axes have been reduced to highlight the optimum area, on average, 305 patients are to be screened for the selection of the top 5% best responders.

4 Discussion

4.1 Opportunities and challenges of combining complex QSP models and trial simulation

PBPK and PD models, QSP models, exposure response modeling and trial simulations are often regarded as distinct modeling paradigms under the umbrella of model-informed drug development⁹⁴. There exist to date only few integrated modeling frameworks which couple QSP models and trial simulations^{18;95–97}, most probably because they are complex to set up, multiscale by design, require a multi-stakeholder engagement and considerable effort for verification and validation^{98;99}. One of the main hurdles for complex models (e.g. in QSP) is the high-dimensional parameter space which needs to be informed or calibrated. High complexity can be a necessary feature¹⁰⁰ of the model and can also increase its robustness: different biological contexts can be covered and thus heterogeneous data from multiple sources can be used for calibration. On the other hand, high-dimensional models are prone to identifiability issues and parameter estimation suffers from non-convex and multi-modal objective functions with gradients that are computationally expensive to evaluate¹⁰¹. To add to this complexity, there is no consensus on the best method, and performance of such methods are model-specific¹⁰¹. We have therefore chosen an approach that minimized the number of parameters to calibrate. We have coupled our disease model, our SoC treatments models, and the treatment model of the immunomodulator OM-85²⁹ into a QSP model. Central to the disease model setup has been the choice of an equilibrium assumption (i.e. defining the system in terms of stationary dynamics), which allowed for calculating rather than calibrating a large fraction of the free parameters. In this way, from the set of 139 parameters present in the final QSP model, 70 could be constrained by the equilibrium approach; with 46 others being fixed with information from the literature, effectively leaving only 23 unknown parameters for calibration (i.e. with SoC data). The drawback of this method is that the pathogenesis, natural evolution and resolution of the disease are not described and that the used datasets (here skin proteomics in pediatric patients) convey context specificity (e.g. the model would need to be re-informed for adults and serum biomarkers).

While such a QSP model can resolve several questions around patient profiles and dosing-regimen, simulation studies of entire clinical trials needs a dedicated workflow. We therefore set up a virtual population of AD patients to reproduce the between-patient variability of our reference clinical trial²⁸. We used arm designs (intervention, dosing-regimen), eligibility criteria and trial duration to perform trial simulations to predict drug efficacy according to a variety of trial protocols, therefore covering a wide spectrum of potential variables for trial design. Classically, in trial simulation used for statistical design, efficacy is considered as a constant by aggregating efficacy across all considered historical trials. However, when data exhibit heterogeneous treatment effects, uncertainty needs to be acknowledged and variability assessed by meta-regression and model-based meta-analysis. In fact, our QSP workflow can be considered as a corner case of model-based meta-analysis where several inputs from patient characteristics and posology are translated through mechanistic relationships into modulation of the efficacy. By using Monte-Carlo trial simulations (sometimes termed microsimulations) for several protocols and by assessing resulting simulation data for sample and effect size, we found a reduced sensitivity of efficacy optimization to dosing regimen (meaning that the dosing scheme from the reference trial is probably already optimal) compared to the choice of a relative severity endpoint over an absolute one which seems to harbor more potential for trial design optimization. In addition, this framework can be readily used for follow-up investigations like recruitment or cost-effectiveness considerations. One of the drawbacks of such trial simulation method (as compared to Monte Carlo simulation with a constant drug effect) is the high computational cost (per virtual patient) and the large storage and memory requirements when simulating many virtual patients.

4.2 Towards the holy grail of model-informed predictive biomarker exploration

Adjusting the dose and the treatment to each patient's individual characteristics are the guiding principles of personalized medicine. Yet, today the gold standard for the approval of new drugs by regulatory agencies are placebo-controlled randomized clinical trials that report efficacy of

entire populations of patients (see our discussion in Courcelles et al.¹⁰² from a health technology assessment perspective). A way to better individualize drug development could be to use patient stratification and subgroup analysis, optimally performed on the basis of easily measurable biomarkers as objective (and ideally validated) predictors of treatment effect (predictive biomarkers). In fact, trials using biomarkers have an almost doubled overall probability of success compared to trials without biomarkers, notably in Phase I and II according to an analysis of Wong et al.¹⁰³, but the use of biomarkers outside of oncology is not very frequent¹⁰³. To be implemented early-on, exploration of potential biomarkers has to be started prior to Phase I trials so that validation and qualification of biomarkers can be tackled during clinical development¹⁰⁴. While in targeted oncology, markers (*i.e.* mutations) indicating the diseases vulnerability to treatment modulation are straightforward in their interpretation, inflammatory diseases often involve an intricate network of immune signaling and cellular pathways so that the search for biomarkers signatures is like a needle in the haystack⁸⁵. In AD, progress has been made to identify biomarkers of severity such as systemic levels of the chemokine C-C motif ligand 17/thymus, an activation-regulated chemokine and chemoattractant of Th2 cells, which shows robust correlation with AD clinical severity, at both baseline and during therapy and thus could be used as an objective surrogate for treatment effect¹⁰⁵, but if such biomarker can also predict treatment effect for a specific investigational drug of interest is still an open question.

The motivation behind this study was to use our QSP workflow to map the body of evidence of the relationships between biomarker levels and disease severity, in line with the immune dysregulation system involved in AD, onto a reference trial so that synthetic trial data is enriched with such biologically plausible information. As a proof of concept and to increase the practical applicability of our workflow, we tested several approaches to identify biomarkers predictive of treatment effect. We identified that neither a linearized surrogate model approach nor subgroup analysis by severity only lead to sufficient stratification potential, which a) plausibly underlines the difficulty to find predictive biomarkers in the context of AD (clinical data are in fact often analyzed using these statistical techniques), and b) emphasizes the need for methods which take into account non-linear effects for this purpose. Only an input dimensionality reduction approach, based on global sensitivity

analysis proved accurate enough for identification of the skin biomarkers most predictive of OM-85 treatment effect. We did not intend to oversimplify the structural model (e.g. by lumping techniques) in view of the anticipated unbiased exploratory nature of the model. As linearization of the model was unsuccessful, the only viable option for simplification was to reduce the number of predictive input baseline biomarkers characterizing each virtual patient by replacing variability by fixed reference values. Quantification of how the input dimensionality reduction approach performs for predicting individual treatment effect makes us confident that embedding it within the *in silico* clinical trial framework can yield a treatment-specific tool that captures the state-of-the-art knowledge about how skin biomarkers are involved in the mechanism of action. As a perspective, this could a) inform the inclusion into the sample collection and analysis plan of e.g. Phase II trials, b) be prospectively validated through comparison with the generated clinical data and c) once qualified could support the clinical evidence by enrichment in sparsely sampled regions of biomarker space. For the use-case of OM-85, we found a subset of four skin biomarkers (alarmins, IL-22, type-2 and regulatory cytokines) predictive of treatment effect and we simulated that patient selection based on this subset may lead to larger effect size. In line with our results and among the skin biomarkers the model identified to be predictive of OM-85 effect, IL-22 and IL-13 expression levels in the skin tissues have been previously identified to strongly and significantly correlate with clinical therapeutic effect in AD¹⁰⁵. Development of such computational approaches goes well in hand with promoting clinical trials designed to validate or reject the predictions made by the model and which could then either be used to improve subsequent (e.g. Phase IIb and III) clinical trial designs or in a learn and confirm paradigm to refine our model¹⁰⁵.

Competing interests statement

N.G., S.A., I.F., T.G., S.M.B., D.L., Y.W., L.E., E.J., C.M., J.B., and A.K. are employees of Novadiscovery. C.S., C.P., and L.L are employees of OM Pharma. Novadiscovery and OM Pharma funded the study.

Acknowledgments

We acknowledge Valentin Vierge and Daniel Šmít for contributing to model building and calibration, Germà Gómez for statistical support, and Louis Philippe for technical support.

Authors contributions

A.K., S.A., L.L., C.P and C.S. supervised the study. N.G., S.A. and A.K. designed the model. N.G, S.A, I.F., T.G., S.M.B., D.L., Y.W., L.E. developed the model. N.G. performed the simulations and analyzed the results. N.G., S.A. and A.K wrote the manuscript. All authors contributed to the discussion of the results and reviewed the manuscript.

Code and data availability

All information needed to reproduce the simulations are described in the references of this manuscript, and the supplementary information. Access to the computational model code, documentation and simulation results are available on Jinko.ai platform upon request to the corresponding author. R scripts developed for analysis of simulation results and producing the figures are also available upon reasonable request.

Supplementary Information

The Supplementary File contains: (S1) Supplementary Methods where the (S1.1) Model, (S1.2) Calibration, (S1.3 Biomarker selection, and (S1.4) Bootstrapping procedure are described in details; (S2) Supplementary Results for (S2.1) Assessment of the efficacy sensitivity to the TCS related trial protocol settings, (S2.2) Performance baseline AD severity as a marker of treatment effect, and (S2.3) Identification of predictive biomarkers using a surrogate model approach; (S3) Supplementary Figures and (S4) Supplementary Tables.

References

- [1] S. Barbarot, S. Auziere, A. Gadkari, G. Girolomoni, L. Puig, E. L. Simpson, D. J. Margolis, M. Bruin-Weller, and L. Eckert. Epidemiology of atopic dermatitis in adults: Results from an international survey. *Allergy*, 73(6):1284–1293, February 2018. doi: 10.1111/all.13401.
- [2] Jonathan I. Silverberg, Joel M. Gelfand, David J. Margolis, Mark Boguniewicz, Luz Fonacier, Mitchell H. Grayson, Eric L. Simpson, Peck Y. Ong, and Zelma C. Chiesa Fuxench. Patient burden and quality of life in atopic dermatitis in US adults. *Annals of Allergy, Asthma & Immunology*, 121(3):340–347, September 2018. doi: 10.1016/j.anai.2018.07.006.
- [3] Eric L. Simpson, Thomas Bieber, Laurent Eckert, Richard Wu, Marius Ardeleanu, Neil M.H. Graham, Gianluca Pirozzi, and Vera Mastey. Patient burden of moderate to severe atopic dermatitis (AD): Insights from a phase 2b clinical trial of dupilumab in adults. *Journal of the American Academy of Dermatology*, 74(3):491–498, March 2016. doi: 10.1016/j.jaad.2015.10.043.
- [4] Torsten Zuberbier, Seth J. Orlow, Amy S. Paller, Alain Taïeb, Roger Allen, José M. Hernanz-Hermosa, Jorge Ocampo-Candiani, Margaret Cox, Joanne Langeraar, and Jan C. Simon. Patient perspectives on the management of atopic dermatitis. *Journal of Allergy and Clinical Immunology*, 118(1):226–232, July 2006. doi: 10.1016/j.jaci.2006.02.031.
- [5] Daniel C. Butler, Eric Simpson, Emma Guttman-Yassky, Lawrence F. Eichenfield, Alexandra K. Golant, John Y. M. Koo, April W. Armstrong, Andrew F. Alexis, Peter A. Lio, Justin W. Marson, and Mark Lebwohl. The atopic dermatitis spectrum disorder. recognizing the clinical heterogeneity in patients with atopic related skin conditions in order to improve therapeutic decision-making and outcomes: an expert panel consensus statement. *Journal of Dermatological Treatment*, 33(4):2397–2399, August 2021. doi: 10.1080/09546634.2021.1966356.
- [6] Sinéad M Langan, Alan D Irvine, and Stephan Weidinger. Atopic dermatitis. *The Lancet*, 396(10247):345–360, August 2020. doi: 10.1016/s0140-6736(20)31286-1.
- [7] Kunal Malik, Kerry D. Heitmiller, and Tali Czarnowicki. An update on the pathophysiology of atopic dermatitis. *Dermatologic Clinics*, 35(3):317–326, July 2017. doi: 10.1016/j.det.

2017.02.006.

- [8] Thomas Bieber. Atopic dermatitis: an expanding therapeutic pipeline for a complex disease. *Nature Reviews Drug Discovery*, 21(1):21–40, August 2021. doi: 10.1038/s41573-021-00266-6.
- [9] Paola Facheris, Jane Jeffery, Ester Del Duca, and Emma Guttman-Yassky. The translational revolution in atopic dermatitis: the paradigm shift from pathogenesis to treatment. *Cellular & Molecular Immunology*, March 2023. doi: 10.1038/s41423-023-00992-4. URL <https://doi.org/10.1038/s41423-023-00992-4>.
- [10] Quinn Thibodeaux, Mary Patricia Smith, Karen Ly, Kristen Beck, Wilson Liao, and Tina Bhutani. A review of dupilumab in the treatment of atopic diseases. *Human Vaccines & Immunotherapeutics*, 15(9):2129–2139, March 2019. doi: 10.1080/21645515.2019.1582403.
- [11] Thomas Bieber, Claudia Traidl-Hoffmann, Georg Schäppi, Roger Lauener, Cezmi Akdis, and Peter Schmid-Grendlmeier. Unraveling the complexity of atopic dermatitis: The CK-CARE approach toward precision medicine. *Allergy*, 75(11):2936–2938, February 2020. doi: 10.1111/all.14194.
- [12] Th. Bieber. Atopic dermatitis 2.0: from the clinical phenotype to the molecular taxonomy and stratified medicine. *Allergy*, pages n/a–n/a, October 2012. doi: 10.1111/all.12049.
- [13] S. E. Baron, S. N. Cohen, and C. B. Archer and. Guidance on the diagnosis and clinical management of atopic eczema. *Clinical and Experimental Dermatology*, 37:7–12, April 2012. doi: 10.1111/j.1365-2230.2012.04336.x.
- [14] Peter A. Lio, Margaret Lee, Jennifer LeBovidge, Karol G. Timmons, and Lynda Schneider. Clinical management of atopic dermatitis: Practical highlights and updates from the atopic dermatitis practice parameter 2012. *The Journal of Allergy and Clinical Immunology: In Practice*, 2(4):361–369, July 2014. doi: 10.1016/j.jaip.2014.02.015.
- [15] Jonathan I. Silverberg, Eric L. Simpson, April W. Armstrong, Marjolein S. de Bruin-Weller, Alan D. Irvine, and Kristian Reich. Expert perspectives on key parameters that impact interpretation of randomized clinical trials in moderate-to-severe atopic dermatitis. *American Journal*

of *Clinical Dermatology*, 23(1):1–11, October 2021. doi: 10.1007/s40257-021-00639-y.

- [16] Charles F. Schuler, Allison C. Billi, Emanuel Maverakis, Lam C. Tsoi, and Johann E. Gudjonsson. Novel insights into atopic dermatitis. *Journal of Allergy and Clinical Immunology*, 151(5):1145–1154, May 2023. doi: 10.1016/j.jaci.2022.10.023. URL <https://doi.org/10.1016/j.jaci.2022.10.023>.
- [17] Kilian Eyerich, Sara J Brown, Bethany E Perez White, Reiko J Tanaka, Robert Bissonette, Sandipan Dhar, Thomas Bieber, Dirk J Hijnen, Emma Guttman-Yassky, Alan Irvine, et al. Human and computational models of atopic dermatitis: A review and perspectives by an expert panel of the international eczema council. *Journal of Allergy and Clinical Immunology*, 143(1):36–45, 2019. doi: 10.1016/j.jaci.2018.10.033.
- [18] Marco Viceconti, Adriano Henney, and Edwin Morley-Fletcher. In silico clinical trials: how computer simulation will transform the biomedical industry. *International Journal of Clinical Trials*, 3(2):37, May 2016. doi: 10.18203/2349-3259.ijct20161408.
- [19] Leia Wedlund and Joseph Kvedar. Simulated trials: in silico approach adds depth and nuance to the RCT gold-standard. *npj Digital Medicine*, 4(1), August 2021. doi: 10.1038/s41746-021-00492-7.
- [20] Vincent Lemaire, David Bassen, Mike Reed, Roy Song, Samira Khalili, Yi Ting (Kayla) Lien, Lu Huang, Aman P. Singh, Spyros Stamatelos, Dean Bottino, and Fei Hua. From cold to hot: Changing perceptions and future opportunities for quantitative systems pharmacology modeling in cancer immunotherapy. *Clinical Pharmacology & Therapeutics*, 113(5):963–972, November 2022. doi: 10.1002/cpt.2770. URL <https://doi.org/10.1002/cpt.2770>.
- [21] Elisa Domínguez-Hüttinger, Panayiotis Christodoulides, Kosuke Miyauchi, Alan D Irvine, Mariko Okada-Hatakeyama, Masato Kubo, and Reiko J Tanaka. Mathematical modeling of atopic dermatitis reveals double-switch mechanisms underlying 4 common disease phenotypes. *Journal of Allergy and Clinical Immunology*, 139(6):1861–1872, 2017. doi: 10.1016/j.jaci.2016.10.026.

- [22] Takuya Miyano, Alan D Irvine, and Reiko J Tanaka. A mathematical model to identify optimal combinations of drug targets for dupilumab poor responders in atopic dermatitis. *Allergy*, 77(2):582–594, 2022. doi: 10.1101/2021.02.08.21251317.
- [23] Jane P. F. Bai, Justin C. Earp, Jeffry Florian, Rajanikanth Madabushi, David G. Strauss, Yaning Wang, and Hao Zhu. Quantitative systems pharmacology: Landscape analysis of regulatory submissions to the US food and drug administration. *CPT: Pharmacometrics & Systems Pharmacology*, 10(12):1479–1484, November 2021. doi: 10.1002/psp4.12709. URL <https://doi.org/10.1002/psp4.12709>.
- [24] Takuya Miyano, Alan D. Irvine, and Reiko J. Tanaka. Model-based meta-analysis to optimize staphylococcus aureustargeted therapies for atopic dermatitis. *JID Innovations*, 2(3):100110, May 2022. doi: 10.1016/j.xjidi.2022.100110. URL <https://doi.org/10.1016/j.xjidi.2022.100110>.
- [25] M Huber, H Mossmann, and W G Bessler. Th1-orientated immunological properties of the bacterial extract OM-85-BV. *Eur. J. Med. Res.*, 10(5):209–217, May 2005.
- [26] Agnieszka Kaczynska, Martyna Klosinska, Kamil Janeczek, Michał Zarobkiewicz, and Andrzej Emeryk. Promising immunomodulatory effects of bacterial lysates in allergic diseases. *Frontiers in Immunology*, 13, June 2022. doi: 10.3389/fimmu.2022.907149.
- [27] Urs B. Schaad. OM-85 BV, an immunostimulant in pediatric recurrent respiratory tract infections: a systematic review. *World Journal of Pediatrics*, 6(1):5–12, February 2010. doi: 10.1007/s12519-010-0001-x.
- [28] Christine Bodemer, Gerard Guillet, Frederic Cambazard, Franck Boralevi, Stefania Bal-larini, Christian Milliet, Paola Bertuccio, Carlo La Vecchia, Jean-Francois Bach, and Yves de Prost. Adjuvant treatment with the bacterial lysate (om-85) improves manage-ment of atopic dermatitis: A randomized study. *PLoS One*, 12(3):e0161555, 2017. doi: 10.1371/journal.pone.0161555.
- [29] Simon Arsène, Claire Couty, Igor Faddeenkov, Natacha Go, Solène Granjeon-Noriot, Daniel Šmít, Riad Kahoul, Ben Illigens, Jean-Pierre Boissel, Aude Chevalier, Lorenz Lehr, Christian

Pasquali, and Alexander Kulesza. Modeling the disruption of respiratory disease clinical trials by non-pharmaceutical COVID-19 interventions. *Nature Communications*, 13(1), April 2022. doi: 10.1038/s41467-022-29534-8.

[30] Magdalena Trzeciak, Monika Sakowicz-Burkiewicz, Martyna Wesserling, Daria Dobaczewska, Jolanta Glen, Roman Nowicki, and Tadeusz Pawelczyk. Expression of cornified envelope proteins in skin and its relationship with atopic dermatitis phenotype. *Acta dermato-venereologica*, 97(1):36–41, 2017. doi: 10.2340/00015555-2482.

[31] Zheng Sun, Shi Huang, Pengfei Zhu, Feng Yue, Helen Zhao, Ming Yang, Yueqing Niu, Gongchao Jing, Xiaoquan Su, Huiying Li, et al. A microbiome-based index for assessing skin health and treatment effects for atopic dermatitis in children. *mSystems*, 4(4):e00293–19, 2019. doi: 10.1128/mSystems.00293-19.

[32] Donald Y. M. Leung, Agustin Calatroni, Livia S. Zaramela, Petra K. LeBeau, Nathan Dyjack, Kanwaljit Brar, Gloria David, Keli Johnson, Susan Leung, Marco Ramirez-Gama, Bo Liang, Cydney Rios, Michael T. Montgomery, Brittany N. Richers, Clifton F. Hall, Kathryn A. Norquest, John Jung, Irina Bronova, Simion Kreimer, C. Conover Talbot, Debra Crumrine, Robert N. Cole, Peter Elias, Karsten Zengler, Max A. Seibold, Evgeny Berdyshev, and Elena Goleva. The nonlesional skin surface distinguishes atopic dermatitis with food allergy as a unique endotype. *Science Translational Medicine*, 11(480):eaav2685, February 2019. doi: 10.1126/scitranslmed.aav2685.

[33] Taras Lyubchenko, Hannah K. Collins, Elena Goleva, and Donald Y.M. Leung. Skin tape sampling technique identifies proinflammatory cytokines in atopic dermatitis skin. *Annals of Allergy, Asthma & Immunology*, 126(1):46–53.e2, January 2020. doi: 10.1016/j.anai.2020.08.397.

[34] E. Guttman-Yassky, A. Diaz, A. B. Pavel, M. Fernandes, R. Lefferdink, T. Erickson, T. Canter, S. Rangel, X. Peng, R. Li, Y. Estrada, H. Xu, J. G. Krueger, and A. S. Paller. Use of Tape Strips to Detect Immune and Barrier Abnormalities in the Skin of Children With Early-Onset Atopic Dermatitis. *JAMA Dermatol*, 155(12):1358–1370, Dec 2019. doi:

10.1001/jamadermatol.2019.2983.

- [35] Helen He, Robert Bissonnette, Gianni Wu, Aisleen Diaz, Etienne Saint-Cyr Proulx, Catherine Maari, Carolyn Jack, Maudeline Louis, Yeriel Estrada, James G. Krueger, Ning Zhang, Ana B. Pavel, and Emma Guttman-Yassky. Tape strips detect distinct immune and barrier profiles in atopic dermatitis and psoriasis. *Journal of Allergy and Clinical Immunology*, 147(1):199–212, January 2021. doi: 10.1016/j.jaci.2020.05.048.
- [36] Annalisa Patrizi, Bruno Capitanio, Iria Neri, Federica Giacomini, Jo L. Sinagra, Beatrice Raone, and Enzo Berardesca. A double-blind, randomized, vehicle-controlled clinical study to evaluate the efficacy and safety of MAS063dp (ATOPICLAIR™) in the management of atopic dermatitis in paediatric patients. *Pediatric Allergy and Immunology*, 0(0): 080220191244231–???, February 2008. doi: 10.1111/j.1399-3038.2008.00724.x.
- [37] Jolanta Szczepanowska, Adam Reich, and Jacek C. Szepietowski. Emollients improve treatment results with topical corticosteroids in childhood atopic dermatitis: a randomized comparative study. *Pediatric Allergy and Immunology*, 19(7):614–618, January 2008. doi: 10.1111/j.1399-3038.2007.00706.x.
- [38] Brian S. Kim, Michael D. Howell, Kang Sun, Kim Papp, Adnan Nasir, and Michael E. Kuligowski. Treatment of atopic dermatitis with ruxolitinib cream (JAK1/JAK2 inhibitor) or triamcinolone cream. *Journal of Allergy and Clinical Immunology*, 145(2):572–582, February 2020. doi: 10.1016/j.jaci.2019.08.042.
- [39] TA Luger, M Lahfa, R Fölster-Holst, WP Gulliver, R Allen, S Molloy, N Barbier, C Paul, and JD Bos. Long-term safety and tolerability of pimecrolimus cream 1% and topical corticosteroids in adults with moderate to severe atopic dermatitis. *Journal of Dermatological Treatment*, 15(3):169–178, June 2004. doi: 10.1080/09546630410033781.
- [40] Jogarao V.S. Gobburu and Lawrence J. Lesko. Quantitative disease, drug, and trial models. *Annual Review of Pharmacology and Toxicology*, 49(1):291–301, February 2009. doi: 10.1146/annurev.pharmtox.011008.145613.
- [41] Jean-Louis Palgen, Angélique Perrillat-Mercerot, Nicoletta Ceres, Emmanuel Peyronnet,

Matthieu Coudron, Elliott Tixier, Ben M. W. Illigens, Jim Bosley, Adèle L’Hostis, and Claudio Monteiro. Integration of heterogeneous biological data in multiscale mechanistic model calibration: Application to lung adenocarcinoma. *Acta Biotheoretica*, 70(3), July 2022. doi: 10.1007/s10441-022-09445-3.

[42] Peter L. Bonate. *Pharmacokinetic-Pharmacodynamic Modeling and Simulation*. Springer US, 2011. doi: 10.1007/978-1-4419-9485-1.

[43] Kilian Eyerich and Stefanie Eyerich. Th22 cells in allergic disease. *Allergo journal international*, 24(1):1–7, 2015. doi: 10.1007/s40629-015-0039-3.

[44] Jung Eun Kim, Jong Sic Kim, Dae Ho Cho, and Hyun Jeong Park. Molecular mechanisms of cutaneous inflammatory disorder: atopic dermatitis. *International journal of molecular sciences*, 17(8):1234, 2016. doi: 10.3390/ijms17081234.

[45] Gregory F Sonnenberg, Lynette A Fouser, and David Artis. Border patrol: regulation of immunity, inflammation and tissue homeostasis at barrier surfaces by il-22. *Nature immunology*, 12(5):383–390, 2011. doi: 10.1038/ni.2025.

[46] A Konur, U Schulz, G Eissner, Reinhard Andreesen, and E Holler. Interferon (ifn)- γ is a main mediator of keratinocyte (hacat) apoptosis and contributes to autocrine ifn- γ and tumour necrosis factor- α production. *British Journal of Dermatology*, 152(6):1134–1142, 2005. doi: 10.1111/j.1365-2133.2005.06508.x.

[47] Takashi Nomura, Tetsuya Honda, and Kenji Kabashima. Multipolarity of cytokine axes in the pathogenesis of atopic dermatitis in terms of age, race, species, disease stage and biomarkers. *International immunology*, 30(9):419–428, 2018. doi: 10.1093/intimm/dxy015.

[48] Patrick M. Brunner, Emma Guttman-Yassky, and Donald Y.M. Leung. The immunology of atopic dermatitis and its reversibility with broad-spectrum and targeted therapies. *Journal of Allergy and Clinical Immunology*, 139(4):S65–S76, April 2017. doi: 10.1016/j.jaci.2017.01.011.

[49] A Halawi, O Abbas, and M Mahalingam. S100 proteins and the skin: a review. *Journal of the European Academy of Dermatology and Venereology*, 28(4):405–414, 2014. doi:

10.1111/jdv.12237.

- [50] Robert Kantor and Jonathan I Silverberg. Environmental risk factors and their role in the management of atopic dermatitis. *Expert review of clinical immunology*, 13(1):15–26, 2017. doi: 10.1080/1744666x.2016.1212660.
- [51] Joshua L Owen, Paras P Vakharia, and Jonathan I Silverberg. The role and diagnosis of allergic contact dermatitis in patients with atopic dermatitis. *American journal of clinical dermatology*, 19(3):293–302, 2018. doi: 10.1007/s40257-017-0340-7.
- [52] Wendy F. Davidson, Donald Y.M. Leung, Lisa A. Beck, Cecilia M. Berin, Mark Boguniewicz, William W. Busse, Talal A. Chatila, Raif S. Geha, James E. Gern, Emma Guttman-Yassky, Alan D. Irvine, Brian S. Kim, Heidi H. Kong, Gideon Lack, Kari C. Nadeau, Julie Schwaninger, Angela Simpson, Eric L. Simpson, Jonathan M. Spergel, Alkis Togias, Ulrich Wahn, Robert A. Wood, Judith A. Woodfolk, Steven F. Ziegler, and Marshall Plaut. Report from the national institute of allergy and infectious diseases workshop on “atopic dermatitis and the atopic march: Mechanisms and interventions”. *Journal of Allergy and Clinical Immunology*, 143(3):894–913, March 2019. doi: 10.1016/j.jaci.2019.01.003.
- [53] Natalija Novak and Thomas Bieber. The role of dendritic cell subtypes in the pathophysiology of atopic dermatitis. *Journal of the American Academy of Dermatology*, 53(2):S171–S176, 2005. doi: 10.1016/j.jaad.2005.04.060.
- [54] Pawinee Rerknimitr, Atsushi Otsuka, Chisa Nakashima, and Kenji Kabashima. The etiopathogenesis of atopic dermatitis: barrier disruption, immunological derangement, and pruritus. *Inflammation and regeneration*, 37(1):1–15, 2017. doi: 10.1186/s41232-017-0044-7.
- [55] Peck Y Ong. New insights in the pathogenesis of atopic dermatitis. *Pediatric research*, 75(1):171–175, 2014. doi: 10.1038/pr.2013.196.
- [56] Thomas Bieber. The pro-and anti-inflammatory properties of human antigen-presenting cells expressing the high affinity receptor for ige (fc ϵ ri). *Immunobiology*, 212(6):499–503, 2007. doi: 10.1016/j.imbio.2007.03.001.
- [57] Sakeen W Kashem, Muzlifah Haniffa, and Daniel H Kaplan. Antigen-presenting

cells in the skin. *Annu Rev Immunol*, 35(1):469–499, 2017. doi: 10.1146/annurev-immunol-051116-052215.

[58] Rachana Agrawal, Julia A Wisniewski, and Judith A Woodfolk. The role of regulatory t cells in atopic dermatitis. *Pathogenesis and Management of Atopic Dermatitis*, 41:112–124, 2011. doi: 10.1159/000323305.

[59] J Thyssen, M Rinnov, and C Vestergaard. Disease mechanisms in atopic dermatitis: A review of aetiological factors. *Acta Dermato Venereologica*, 100(12):adv00162, 2020. doi: 10.2340/00015555-3512.

[60] KU Schallreuter, Ch Levenig, J Berger, J Umbert, RK Winkelmann, L Wegener, O Correia, O Chosidow, Ph Saiag, S Bastuji-Garin, et al. Severity scoring of atopic dermatitis: the SCORAD index. *Dermatology*, 186(1):23–31, 1993. doi: 10.1159/000247298.

[61] J. M. Hanifin, M. Thurston, M. Omoto, R. Cherill, S. J. Tofte, M. Graeber, and The Easi Evaluator Group. The eczema area and severity index (EASI): assessment of reliability in atopic dermatitis. *Experimental Dermatology*, 10(1):11–18, February 2001. doi: 10.1034/j.1600-0625.2001.100102.x.

[62] Stephan Weidinger and Natalija Novak. Atopic dermatitis. *The Lancet*, 387(10023): 1109–1122, March 2016. doi: 10.1016/s0140-6736(15)00149-x.

[63] Lawrence F. Eichenfield, Wynniss L. Tom, Timothy G. Berger, Alfons Krol, Amy S. Paller, Kathryn Schwarzenberger, James N. Bergman, Sarah L. Chamlin, David E. Cohen, Kevin D. Cooper, Kelly M. Cordoro, Dawn M. Davis, Steven R. Feldman, Jon M. Hanifin, David J. Margolis, Robert A. Silverman, Eric L. Simpson, Hywel C. Williams, Craig A. Elmetts, Julie Block, Christopher G. Harrod, Wendy Smith Begolka, and Robert Sidbury. Guidelines of care for the management of atopic dermatitis. *Journal of the American Academy of Dermatology*, 71(1):116–132, July 2014. doi: 10.1016/j.jaad.2014.03.023.

[64] Peter J. Barnes. Anti-inflammatory actions of glucocorticoids: Molecular mechanisms. *Clinical Science*, 94(6):557–572, June 1998. doi: 10.1042/cs0940557.

[65] N. J. Reynolds and W. I. Al-Daraji. Calcineurin inhibitors and sirolimus: mechanisms of

action and applications in dermatology. *Clinical and Experimental Dermatology*, 27(7): 555–561, October 2002. doi: 10.1046/j.1365-2230.2002.01148.x.

[66] A. Alomar, J. Berth-Jones, J.D. Bos, A. Giannetti, S. Reitamo, T. Ruzicka, J-F. Stalder, and K. Thestrup-Pedersen and. The role of topical calcineurin inhibitors in atopic dermatitis:. *British Journal of Dermatology*, 151(s70):3–27, December 2004. doi: 10.1111/j.1365-2133.2004.06269.x.

[67] Luís Uva, Diana Miguel, Catarina Pinheiro, Joana Antunes, Diogo Cruz, João Ferreira, and Paulo Filipe. Mechanisms of action of topical corticosteroids in psoriasis. *International Journal of Endocrinology*, 2012:1–16, 2012. doi: 10.1155/2012/561018.

[68] Gagandeep Kwatra and Sandip Mukhopadhyay. Topical corticosteroids: Pharmacology. In *A Treatise on Topical Corticosteroids in Dermatology*, pages 11–22. Springer Singapore, May 2017. doi: 10.1007/978-981-10-4609-4_2.

[69] Thomas Hultsch, Alexander Kapp, and Jonathan Spergel. Immunomodulation and safety of topical calcineurin inhibitors for the treatment of atopic dermatitis. *Dermatology*, 211(2): 174–187, 2005. doi: 10.1159/000086739.

[70] Sandipan Dhar, Joly Seth, and Deepak Parikh. Systemic side-effects of topical corticosteroids. *Indian Journal of Dermatology*, 59(5):460, 2014. doi: 10.4103/0019-5154.139874.

[71] Sean Christopher Kearney, Marcin Dziekiewicz, and Wojciech Feleszko. Immunoregulatory and immunostimulatory responses of bacterial lysates in respiratory infections and asthma. *Annals of Allergy, Asthma & Immunology*, 114(5):364–369, 2015. doi: 10.1016/j.anai.2015.02.008.

[72] Clark D. Russell, Stefan A. Unger, Marc Walton, and Jürgen Schwarze. The human immune response to respiratory syncytial virus infection. *Clinical Microbiology Reviews*, 30(2): 481–502, April 2017. doi: 10.1128/cmr.00090-16.

[73] Xianwei Wang, Hui Peng, and Zhigang Tian. Innate lymphoid cell memory. *Cellular & Molecular Immunology*, 16(5):423–429, February 2019. doi: 10.1038/s41423-019-0212-6.

[74] S Navarro, G Cossalter, C Chiavaroli, A Kanda, S Fleury, A Lazzari, J Cazareth, T Sparwasser,

D Dombrowicz, N Glaichenhaus, and V Julia. The oral administration of bacterial extracts prevents asthma via the recruitment of regulatory t cells to the airways. *Mucosal Immunology*, 4(1):53–65, January 2011. doi: 10.1038/mi.2010.51.

[75] D H Strickland, S Judd, J A Thomas, A N Larcombe, P D Sly, and P G Holt. Boosting airway t-regulatory cells by gastrointestinal stimulation as a strategy for asthma control. *Mucosal Immunology*, 4(1):43–52, January 2011. doi: 10.1038/mi.2010.43.

[76] Ran Fu, Jian Li, Hua Zhong, Dehong Yu, Xianping Zeng, Mengxia Deng, Yueqi Sun, Weiping Wen, and Huabin Li. Broncho-vaxom attenuates allergic airway inflammation by restoring GSK3 β -related t regulatory cell insufficiency. *PLoS ONE*, 9(3):e92912, March 2014. doi: 10.1371/journal.pone.0092912.

[77] B. Emmerich, K. Pachmann, D. Milatovic, and H.P. Emslander. Influence of OM-85 BV on different humoral and cellular immune defense mechanisms of the respiratory tract. *Respiration*, 59(3):19–23, 1992. doi: 10.1159/000196126.

[78] Mirco Lusuardi, Armando Capelli, Sonia Carli, Elisa L. Spada, Alberto Spinazzi, and Claudio F. Donner. Local airways immune modifications induced by oral bacterial extracts in chronic bronchitis. *Chest*, 103(6):1783–1791, June 1993. doi: 10.1378/chest.103.6.1783.

[79] J R Mora and U H von Andrian. Differentiation and homing of IgA-secreting cells. *Mucosal Immunology*, 1(2):96–109, March 2008. doi: 10.1038/mi.2007.14.

[80] So-Yeon Lee, Eun Lee, Yoon Mee Park, and Soo-Jong Hong. Microbiome in the gut-skin axis in atopic dermatitis. *Allergy, Asthma & Immunology Research*, 10(4):354, 2018. doi: 10.4168/aair.2018.10.4.354.

[81] J P Boissel, M Cucherat, W Li, G Chatellier, F Gueyffier, M Buyse, F Boutitie, P Nony, M Haugh, and G Mignot. The problem of therapeutic efficacy indices. 3. comparison of the indices and their use. *Therapie*, 54(4):405–411, 1999.

[82] Zuozhen Wang. Comparison of sample size by bootstrap and by formulas based on normal distribution assumption. *Therapeutic Innovation & Regulatory Science*, 53(2):170–175, March 2019. doi: 10.1177/2168479018778280.

- [83] Cynthia Huber, Tim Friede, Julia Stingl, and Norbert Benda. Classification of companion diagnostics: A new framework for biomarker-driven patient selection. *Therapeutic Innovation & Regulatory Science*, 56(2):244–254, November 2021. doi: 10.1007/s43441-021-00352-2. URL <https://doi.org/10.1007/s43441-021-00352-2>.
- [84] Daphne Bakker, Marjolein de Bruin-Weller, Julia Drylewicz, Femke van Wijk, and Judith Thijs. Biomarkers in atopic dermatitis. *Journal of Allergy and Clinical Immunology*, February 2023. doi: 10.1016/j.jaci.2023.01.019. URL <https://doi.org/10.1016/j.jaci.2023.01.019>.
- [85] Styliani Mastrafitsi, Georgia Vrioni, Michail Bakakis, Electra Nicolaidou, Dimitrios Rigopoulos, Alexander J. Stratigos, and Stamatis Gregoriou. Atopic dermatitis: Striving for reliable biomarkers. *Journal of Clinical Medicine*, 11(16):4639, August 2022. doi: 10.3390/jcm11164639. URL <https://doi.org/10.3390/jcm11164639>.
- [86] Ich guideline e11a on pediatric extrapolation ema/chmp/ich/205218/2022. URL https://www.ema.europa.eu/en/documents/scientific-guideline/draft-ich-guideline-e11a-pediatric-extrapolation-step-2b_en.pdf.
- [87] D. R. Matthews, J. P. Hosker, A. S. Rudenski, B. A. Naylor, D. F. Treacher, and R. C. Turner. Homeostasis model assessment: insulin resistance and β -cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia*, 28(7):412–419, July 1985. doi: 10.1007/bf00280883. URL <https://doi.org/10.1007/bf00280883>.
- [88] Peter L. Bonate, Amit Desai, Ahsan Rizwan, Zheng Lu, and Stacey Tannenbaum. Nonlinear mixed effects modeling in systems pharmacology. In *Systems Pharmacology and Pharmacodynamics*, pages 255–276. Springer International Publishing, 2016. doi: 10.1007/978-3-319-44534-2_12. URL https://doi.org/10.1007/978-3-319-44534-2_12.
- [89] Wei Chang Colin Tan, Sanjna Nilesh Nerurkar, Hai Yun Cai, Harry Ho Man Ng, Duoduo Wu, Yu Ting Felicia Wee, Jeffrey Chun Tatt Lim, Joe Yeong, and Tony Kiat Hon Lim.

Overview of multiplex immunohistochemistry/immunofluorescence techniques in the era of cancer immunotherapy. *Cancer Communications*, 40(4):135–153, April 2020. doi: 10.1002/cac2.12023.

[90] Tongli Zhang, Ioannis P. Androulakis, Peter Bonate, Limei Cheng, Tomáš Helikar, Jaimit Parikh, Christopher Rackauckas, Kalyanasundaram Subramanian, Carolyn R. Cho, Ioannis P. Androulakis, Peter Bonate, Ivan Borisov, Gordon Broderick, Limei Cheng, Valeriu Damian, Rafael Dariolli, Oleg Demin, Nicholas Ellinwood, Dirk Fey, Abhishek Gulati, Tomas Helikar, Eric Jordie, Cynthia Musante, Jaimit Parikh, Christopher Rackauckas, Julio Saez-Rodriguez, Eric Sobie, Kalyanasundaram Subramanian, and Carolyn R. Cho and. Two heads are better than one: current landscape of integrating QSP and machine learning. *Journal of Pharmacokinetics and Pharmacodynamics*, 49(1):5–18, February 2022. doi: 10.1007/s10928-022-09805-z. URL <https://doi.org/10.1007/s10928-022-09805-z>.

[91] A Saltelli. Sensitivity analysis; saltelli a, chan k, scott m, editors, 2000.

[92] I.M Sobol'. Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *Mathematics and Computers in Simulation*, 55(1-3):271–280, February 2001. doi: 10.1016/s0378-4754(00)00270-6.

[93] Agnieszka Bożek and Adam Reich. Assessment of intra- and inter-rater reliability of three methods for measuring atopic dermatitis severity: EASI, objective SCORAD, and IGA. *Dermatology*, 233(1):16–22, 2017. doi: 10.1159/000472711.

[94] Rajanikanth Madabushi, Paul Seo, Liang Zhao, Million Tegenge, and Hao Zhu. Review: Role of model-informed drug development approaches in the lifecycle of drug development and regulatory decision-making. *Pharmaceutical Research*, 39(8):1669–1680, May 2022. doi: 10.1007/s11095-022-03288-w.

[95] José Ramón Gutiérrez-Casares, Javier Quintero, Guillem Jorba, Valentin Junet, Vicente Martínez, Tamara Pozo-Rubio, Baldomero Oliva, Xavier Daura, José Manuel Mas, and Carmen Montoto. Methods to develop an in silico clinical trial: Computational head-to-head comparison of lisdexamfetamine and methylphenidate. *Frontiers in Psychiatry*, 12, November

2021. doi: 10.3389/fpsy.2021.741170.

[96] Francesco Pappalardo, Giulia Russo, Flora Musuamba Tshinanu, and Marco Viceconti.

In silico clinical trials: concepts and early adoptions. *Briefings in Bioinformatics*, 20(5):

1699–1708, June 2018. doi: 10.1093/bib/bby043.

[97] A. Carlier, A. Vasilevich, M. Marechal, J. de Boer, and L. Geris. In silico clinical

trials for pediatric orphan diseases. *Scientific Reports*, 8(1), February 2018. doi:

10.1038/s41598-018-20737-y.

[98] Marco Viceconti, Francesco Pappalardo, Blanca Rodriguez, Marc Horner, Jeff Bischoff,

and Flora Musuamba Tshinanu. In silico trials: Verification, validation and uncertainty

quantification of predictive models used in the regulatory evaluation of biomedical products.

Methods, 185:120–127, January 2021. doi: 10.1016/j.ymeth.2020.01.011.

[99] Flora T. Musuamba, Ine Skottheim Rusten, Raphaëlle Lesage, Giulia Russo, Roberta Bursi,

Luca Emili, Gaby Wangorsch, Efthymios Manolis, Kristin E. Karlsson, Alexander Kulesza,

Eulalie Courcelles, Jean-Pierre Boissel, Cécile F. Rousseau, Emmanuelle M. Voisin, Rossana

Alessandrello, Nuno Curado, Enrico Dall’ara, Blanca Rodriguez, Francesco Pappalardo, and

Liesbet Geris. Scientific and regulatory evaluation of mechanistic in silico drug and disease

models in drug development: Building model credibility. *CPT: Pharmacometrics & Systems*

Pharmacology, 10(8):804–825, July 2021. doi: 10.1002/psp4.12669.

[100] B Ribba, HP Grimm, B Agoram, MR Davies, K Gadkar, S Niederer, N van Riel, J Timmis,

and PH van der Graaf. Methodologies for quantitative systems pharmacology (QSP) models:

Design and estimation. *CPT: Pharmacometrics & Systems Pharmacology*, 6(8):496–498,

July 2017. doi: 10.1002/psp4.12206.

[101] Eva-Maria Kapfer, Paul Stapor, and Jan Hasenauer. Challenges in the calibration of large-

scale ordinary differential equation models. *IFAC-PapersOnLine*, 52(26):58–64, 2019. doi:

10.1016/j.ifacol.2019.12.236.

[102] Eulalie Courcelles, Jean-Pierre Boissel, Jacques Massol, Ingrid Klingmann, Riad Kahoul,

Marc Hommel, Emmanuel Pham, and Alexander Kulesza. Solving the evidence interpretabil-

ity crisis in health technology assessment: A role for mechanistic models? *Frontiers in Medical Technology*, 4, February 2022. doi: 10.3389/fmedt.2022.810315.

[103] Chi Heem Wong, Kien Wei Siah, and Andrew W Lo. Estimation of clinical trial success rates and related parameters. *Biostatistics*, 20(2):273–286, January 2018. doi: 10.1093/biostatistics/kxx069.

[104] Daniel S. W. Tan, George V. Thomas, Michelle D. Garrett, Udai Banerji, Johann S. de Bono, Stan B. Kaye, and Paul Workman. Biomarker-driven early clinical trials in oncology. *The Cancer Journal*, 15(5):406–420, September 2009. doi: 10.1097/ppo.0b013e3181bd0445.

[105] Yael Renert-Yuval, Jacob P. Thyssen, Robert Bissonnette, Thomas Bieber, Kenji Kabashima, DirkJan Hijnen, and Emma Guttman-Yassky. Biomarkers in atopic dermatitis—a review on behalf of the international eczema council. *Journal of Allergy and Clinical Immunology*, 147(4):1174–1190.e1, April 2021. doi: 10.1016/j.jaci.2021.01.013.