Human immune challenge to assess proof-of-mechanism in drug development

Philip Drennan
St Hilda's College
University of Oxford

A thesis submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Hilary 2026

Acknowledgements

I would like to thank my supervisors, Professors Mark Coles, James Fullerton, and Duncan Richards, for their guidance and insight throughout this DPhil. Our regular conversations have been a highlight over the past four years. I am also grateful to Jane Holmes for her input on the study design in the initial stages of the project. I further thank my funders, the Kennedy Trust for Rheumatology Research and the John Climax Donation, without whose support this work would not have been possible.

This research would not have been possible without the Translational Pharmacology Group at the University of Oxford. As I did not perform the wet-lab studies myself, I am deeply reliant on the data generated by my colleagues—modern science is a team sport. I am particularly grateful to Dr. Roel de Maeyer and Kate Hollett for their experimental expertise and for providing the foundation upon which my analysis was built.

I would also like to acknowledge Professor Stephen Chambers, who played a significant role in shaping my initial entry into a research career at a formative moment.

To my parents, thank you for your continued support and encouragement.

Finally, and most importantly, I thank my wife, Rachel. Completing a DPhil is a challenge in itself, but doing so while welcoming two children into the world is something else entirely. Your unwavering support, patience, and generosity made this work possible.

Abstract

Background: High attrition rates in immunomodulatory drug development, particularly in Phase II clinical trials, are largely driven by failure to demonstrate efficacy at this key development stage. First-principles and observational evidence indicate that establishing Proof-of-Mechanism (PoM) early is essential to support robust Go/No-go decisions. Human Immune Challenge (HIC) models provide an experimental platform to elicit mechanism-relevant Pharmacodynamic (PD) biomarkers in healthy volunteers, thereby bridging the translational gap between preclinical evidence and Phase I Proof-of-Concept (PoC). However, existing HIC models—including those employing the widely used challenge antigen Keyhole Limpet Haemocyanin (KLH)—are characterised by substantial methodological heterogeneity, limited standardisation, and narrowly restricted outcome measures. In parallel, despite rapid growth in the number of druggable immunological targets and clinical-stage Investigational Medicinal Products (IMPs), there remains a lack of validated HIC paradigms capable of eliciting many target pathways in healthy volunteers, precluding mechanism-relevant PD interrogation for a large proportion of immunomodulatory IMPs in Phase I. Collectively, these limitations constrain the feasibility and decision-making value of HIC-enabled PoM studies in contemporary drug development.

This thesis tests the hypothesis that systematic optimisation of challenge agent, dose, adjuvant, tissue, and assessment timing enables HIC models to generate robust, mechanism-specific PD readouts in healthy volunteers. It further hypothesises that formal integration of these readouts within a Bayesian decision-theoretic framework yields more informative and economically optimal Go/No-go decisions than conventional Phase I development strategies that either lack mechanism-relevant PD assessment or incorporate such evidence informally.

Methods: To address these aims, a multi-modal experimental and analytical research programme was conducted:

1. An observational study of Modified Vaccinia Ankara (MVA) vaccination to characterise immunogenicity and assess its potential as a viral HIC model relevant to PD evaluation of immunomodulatory agents targeting innate viral signalling, T helper type 1 (T_h1), and $CD8^+$ T-cell pathways.

2. A systematic review of 46 studies employing KLH challenge in drug development to identify design limitations and opportunities for model optimisation.

3. Two randomised, single-blind HIC studies (KLH1 and KLH2) designed to optimise the KLH model, including evaluation of adjuvants (Aluminium oxyhydroxide 2% aqueous

suspension (Alhydrogel) vs. Montanide ISA-51 (Montanide)), Bayesian modelling of the intradermal KLH rechallenge dose–response relationship using Maximum effect (E_{max}) models, and characterisation of the temporal evolution of cutaneous immune responses via multiparameter flow cytometry.

4. A pilot study evaluating Contrast-Enhanced Ultrasound (CEUS) to improve identification of primary draining lymph nodes–a site of increasing interest for invasive sampling to enable tissue-based readouts of HIC and PD activity.

5. A Bayesian decision-theoretic simulation examining how HIC-derived PD data can be formally integrated with economic utility to optimise Go/No-go decision thresholds.

Results: The Modified Vaccinia Ankara – Bavarian Nordic (MVA-BN) study demonstrated the feasibility of MVA-BN as a viral HIC agent, subject to further optimisation. MVA-BN elicited robust T_h1-biased and $CD8^+$ T-cell responses; however, substantial inter-individual variability was observed in serological and cellular responses measured in blood. These findings indicate a need for further characterisation of statistical operating characteristics, dose–response relationships, and endpoint selection, particularly with respect to immune responses at the site of intradermal administration. In the KLH optimisation studies, Montanide elicited significantly greater systemic T-cell responses (IFNγ and IL-4) than Alhydrogel, despite comparable serological effects, with both adjuvants demonstrating greater immunogenicity than unadjuvanted subunit KLH. Bayesian modelling identified 10 μg as the optimal intradermal rechallenge dose, balancing response consistency with expected modulation-sensitivity. Temporal analyses revealed a shift in the cutaneous immune infiltrate from a mixed myeloid (classical monocyte)–lymphoid phenotype at 48 hours to a T_h1-polarised response with enrichment of intermediate and non-classical monocytes at later timepoints (Days 5 and 14). The CEUS study was terminated early for futility due to failure to reliably identify target lymph node enhancement, demonstrating the utility of adaptive experimental designs in efficiently evaluating feasibility in experimental medicine studies. Finally, the decision-theoretic simulation showed that conventional frequentist significance thresholds are likely suboptimal from an economic perspective, supporting the use of utility-calibrated probabilistic decision rules to guide Go/No-go decisions.

Conclusions: This thesis delivers a standardised and optimised protocol for the KLH HIC model and provides PoC evidence for MVA-BN as a promising viral HIC agent capable of eliciting innate and adaptive antiviral immune pathways. It demonstrates that assessment timepoints must be tailored to the biomarker and mechanism of interest, with

early timepoints (e.g., 48 hours) optimally capturing innate and myeloid responses and later timepoints (Days 5–14) required to interrogate adaptive T-cell modulation and local tissue adaptation in KLH HIC. By integrating optimised experimental platforms with rigorous Bayesian decision-theoretic analysis, this work advances the feasibility and decision-making value of HIC-enabled Phase I PoM assessment in immunomodulatory drug development.

Declarations

I declare that, unless otherwise stated, the work presented in this thesis is my own. I did not personally perform laboratory bench work for the studies described herein. My independent contributions comprised the conception, design, execution, statistical analysis, interpretation, and reporting of all studies included in the thesis.

Several components of the work relied on collaboration, where tissue processing, specialist laboratory assays, and ultrasound imaging were conducted by trained colleagues under my direction or in collaboration. For chapters based on published work, individual contributions are described in the corresponding publications. I am particularly grateful to the following individuals and organisations for their contributions to specific aspects of the work:

- **MVA-BN study (Chapter 2):** Stephanie Harris (tissue processing and ELISpot assays); Roel de Maeyer (panel design, flow cytometry assays, data processing, and gating); Kate Hollett (panel design, flow cytometry assays); Nicholas Provine (transcriptomic analysis - reported in chapter discussion); UK Health Security Agency (Ashley Otter, Scott Jones; serological analyses).

- **KLH systematic review (Chapter 3):** Dimitrios Karpoinis (manuscript screening).

- **KLH human immune challenge studies (Chapter 4):** Roel de Maeyer (panel design, sample processing and assays; flow cytometry data processing, gating, and expert interpretation); Kate Hollett (panel design, sample processing and assays; serological and ELISpot analyses); Patrick Maclean (transcriptomic sample processing, alignment, and quality control).

- **Microbubble study (Chapter 5):** Paul Lyon, Consultant Radiologist, Oxford University Hospitals NHS Foundation Trust (contrast-enhanced ultrasound imaging).

Contents

List of Figures

List of Abbreviations

Alhydrogel Aluminium oxyhydroxide 2% aqueous suspension

BDT Bayesian Decision Theory

CEUS Contrast-Enhanced Ultrasound

cGAS cyclic GMP-AMP synthase

FDR False Discovery Rate

FNA Fine Needle Aspiration

Go/No-go Go/No-go

HIC Human Immune Challenge

IMP Investigational Medicinal Product

ISG Interferon Stimulated Gene

KLH Keyhole Limpet Haemocyanin

LDI Laser Doppler Imaging

Montanide Montanide ISA-51

MVA Modified Vaccinia Ankara

MVA-BN Modified Vaccinia Ankara – Bavarian Nordic

NU Net Expected Utility

OPXV Orthopoxvirus

PD Pharmacodynamic

PoC Proof-of-Concept

PoM Proof-of-Mechanism

R&D Research and Development

SHELF Sheffield Elicitation Framework

SLE Systemic Lupus Erythematosus

STING Stimulator of Interferon Genes

Chapter 1

Background

Chapter summary

In this chapter, I outline the motivation for the thesis: immunomodulatory drug development is protracted, expensive, and characterised by high failure rates. I show that establishing PoM at the completion of Phase I clinical trials is a critical determinant of subsequent success during Phase II. I argue that because HIC models enable the elicitation of PD biomarkers in healthy volunteers, they expand the pool of drugs for which PoM assessment is feasible in Phase I studies. This discussion leads directly to the rationale for this DPhil project, which aims to develop and validate fit-for-purpose HIC/PoM paradigms.

1.1 The Pharmaceutical Research and Development (R&D) Productivity Crisis

Pharmaceutical R&D is a high-stakes endeavour where the vast majority of candidates fail to reach the patient. Despite sustained growth in investment, approximately 90% of IMP fail prior to regulatory approval. [1, 2] Although recent longitudinal data from 2006–2022 indicate a modest recovery in overall industry success rates [3], the attrition remains severe,

particularly for novel mechanisms. The progressive escalation in resource commitment required to advance an immunomodulatory candidate is summarised in Table 1.1. Notably, the probability that an agent entering Phase I ultimately achieves regulatory approval is only 15.1%, meaning that the substantial investment in early clinical development carries an exceptionally high risk of failure.

This high attrition drives the exorbitant cost of drug development, currently estimated at >USD 2.7 billion per approved agent. [4] Importantly, the locus of this failure has shifted in recent years. Whereas safety signals once drove attrition, failure of immunomodulatory IMPs is now predominantly (>80%) attributable to a lack of demonstrated efficacy in Phase II clinical trials. [5]

These figures stand in stark contrast to impressive recent technological advances in molecular medicine, high-throughput technologies, and rational drug design. This paradox—declining output despite improving technology—is concisely summarised as *Eroom's Law* (an inversion of Moore's Law). Defined by Scannell et al., it describes the trend where the number of new drugs approved per billion US dollars spent on R&D has halved approximately every nine years since 1950. [8] The urgency of this efficiency crisis is further underscored by a recent analysis of trial terminations between 2013 and 2023. Bowling et al. report that the number of Phase II and III trial terminations has doubled over the last decade, identifying two critical drivers beyond simple efficacy failure. [9] First, 'Strategic/Business' terminations have increased, reflecting an increasingly risk-averse environment where assets lacking robust early validation are vulnerable to portfolio reprioritisation. Second, terminations due to enrolment challenges have risen, particularly in Phase II, signalling a practical bottleneck in the global clinical infrastructure. Consequently, the solution to the productivity crisis cannot simply be to increase the volume of candidates entering Phase II,

Table 1.1: Characteristics of clinical development phases for immunomodulatory drugs. Descriptions of objectives and participant pools are derived from Food and Drug Administration guidance. [6] Costs represent estimated per-study operational expenditures, while the final row reflects the total cost per approved New Molecular Entity. Probability of success indicates the likelihood of an agent at the start of that phase eventually achieving regulatory approval (Autoimmune/Inflammation indication). Estimates derived from published industry data. [4, 7]

Phase	Primary Objectives	Participants	Prob. of Approval	Est. Cost
Phase I	**Safety, Tolerability & Dosing.** Tests behaviour of the IMP in human to define Pharmacokinetic and common acute adverse effects. PoM may be assessed in specialised "Phase Ib" cohorts, though this is not routine.	20–80 Healthy volunteers*	15.1%	$4–10M
Phase II	**Method Refinement & Safety.** Administration to patients to gather further safety data and refine research protocols for Phase III. While often used for efficacy signals (PoC, cohorts typically too small to definitively prove benefit.	Several Hundred Patients	21.2%	$15–25M
Phase III	**Treatment Benefit ('Pivotal').** Large-scale confirmation of whether the IMP offers a treatment benefit to a specific population. Longer duration allows detection of rarer or long-term side effects.	300–3,000 Patients	63.7%	$50–100M+
Submission	**Marketing Authorisation.** Regulatory review of NDA/BLA filing. Phase IV studies may be performed post-marketing approval to assess real world efficacy and safety	N/A	80.3%	**$2.7B****

*Patients are recruited instead of healthy volunteers if the drug has high toxicity (e.g., antineoplastic agents, cell and gene therapies)
**Total approximate cost per approved new drug (accounting for all failures).

as the clinical machinery cannot accommodate the throughput. Instead, development programs require improved decision-making tools to ensure that only assets with the highest probability of success consume these limited operational resources.

1.2 Bridging the Translational Divide: Proof-of-Mechanism

To address the efficacy gap described by Cook et al., the industry has shifted focus from simple safety and pharmacokinetic assessment to validating the biological hypothesis early in development.[5] This strategic pivot is exemplified by Pfizer's *Three Pillars* framework originally reported by Morgan et al., which stipulates that a candidate must demonstrate: (1) exposure at the site of action; (2) binding to the pharmacological target; and (3) expression of pharmacological activity (Proof-of-Mechanism or PoM). Recent analyses of Pfizer's portfolio by Wu et al. confirmed that New Molecular Entities (NMEs) achieving a positive Three-Pillar rating—specifically confirming Pillar 3—had a Phase II success rate of 29%, compared to 0% for those with negative ratings. [10] In the same publication, Wu et al. reported an improvement in Pfizer's Phase II success rates from 19% to 53%. This 'turnaround' in R&D efficiency attributed to a suite of measures including adherence to Three-Pillar-based decision structures, providing empirical observational evidence to support the first-principles rationality of the framework.

1.3 Human Immune Challenge Models

Despite the great promise of PoM-informed early phase decision making, substantial challenges prevent widespread application of these approaches in immunomodulatory drug development. Healthy volunteers (the typical population recruited for phase I studies in the

majority of drug-development fields, excluding Oncology) by definition lack the active immune/inflammatory pathways targeted by the investigational medicinal product (IMP).

One approach to this problem is to elicit these pathways via controlled exogenous stimulation, an approach referred to in this thesis as Human Immune Challenge (HIC). In HIC, exogenous stimulants are administered to healthy volunteers to elicit the activation of pathways, cell populations, and genes which are otherwise quiescent during homeostasis.

By inducing a controlled, disease-relevant phenotype in healthy subjects, HIC permits the rapid and economical assessment of Investigational Medicinal Product (IMP) in early-phase trials. Effective PoM determination via HIC can de-risk development, preventing strategic termination mid-phase by providing reassurance of an increased probability of success or prompting a decision to terminate early, before costly late phase investment. [11, 12]

1.3.1 Rationale and Utility

Experimental medicine immune challenge paradigms represent an under-utilised means to address the 'translational gap' between preclinical models and patient trials. While animal models are essential for safety, they frequently fail to predict human efficacy due to species-specific divergence in immune signalling. [13, 14] Conversely, *ex vivo* stimulation assays often lack the systemic complexity required to model drug distribution and tissue pharmacodynamics.

HIC bridges this divide by enabling controlled investigation of human immunology within a strictly controlled setting. Unlike observational patient studies, HIC allows for the standardisation of key experimental variables, including:

- **Inclusion / Exclusion:** Optimising variability (due to baseline demographic or

disease-related heterogeneity) versus generalisability, or explicitly recruiting populations with specific characteristics relevant to the IMP development programme, such as where the target indication is predominantly associated with particular demographic groups.

- **Primary Dose of Challenge:** Ensuring uniform immune stimulation across cohorts.
- **Timing of Challenge:** Controlling the timing of drug exposure relative to exposure of the challenge agent to ensure the PD signal elicited can be subjected to experimental modulation by the IMP at a biologically relevant timepoint.
- **Complex Endpoints:** Facilitating access to druggable targets and pharmacodynamic endpoints in accessible tissues (e.g., skin, lymph nodes) that are often ethically difficult to sample in patient populations, and enabling arbitrary sampling frequency within ethical and practical constraints.

Such standardisation is critical for rational study design, especially where sample size considerations are based on prior knowledge of expected variability of the primary endpoint.

Broadly, these models can be categorised by the nature of the immune response (Innate vs. Adaptive) and the volunteer's exposure history (Recall vs. Neoantigen). While innate challenges (e.g., LPS, TLR7 agonists) offer rapid readouts, adaptive challenges are critical for assessing drugs targeting T-cell and B-cell interactions—the mechanism of action for many modern biologics.

1.3.2 Limitations of Existing HIC Models

Despite their theoretical utility, the value of HIC is predicated on the existence of standardised techniques with known, disease-relevant response characteristics. Currently, for many

paradigms, no widely accepted or validated protocols exist. This lack of standardisation drives irreproducibility and impairs decision-making in early-phase drug development. [15]

Key barriers preventing the wider adoption of HIC include:

- **Methodological Heterogeneity:** Variations in challenge agents (preparation, purity) and dosing schedules make benchmarking between studies nearly impossible.
- **Lack of Quantitative Rigour:** Many academic HIC studies suffer from small sample sizes (n), flexibility in endpoint selection (potential for selective emphasis and reporting of endpoints post hoc), and a lack of transparency in analytical approaches, leading to noisy data with poor signal-to-noise ratios.
- **Underexplored Endpoints:** While blood biomarkers are commonly assessed, tissue-based readouts—such as lymph node Fine Needle Aspiration (FNA) or skin biopsies—remain underexplored. These sites are often the true intended site of action for immunomodulators, or have the potential to enable more precise readouts of challenge and IMP response, yet rigorous frameworks for their analysis are lacking.
- **Assay Complexity:** The tension between academic research (using complex, bespoke assays) and industry requirements (needing validated, QC-compliant assays) often stalls the translation of these models into regulatory-grade tools.

1.3.3 The Necessity of Tissue-Level Pharmacodynamics

A fundamental limitation of many historical HIC paradigms is their reliance on easily accessible peripheral blood-derived biomarkers as the primary endpoint, for example anti-KLH IgG following KLH challenge. While assessing serum antibody titres is sufficient for evaluating gross immunogenicity or toxicological immunosuppression, it lacks the specificity

required for robust PoM assessment. This inadequacy stems from the fact that circulating antibody levels represent the cumulative output of complex, multi-stage immunological processes. Consequently, by the time a drug effect manifests in serum titres, the signal is often diluted by systemic redundancy that masks subtle effects on specific upstream pathways. Furthermore, the signal is delayed, as seroconversion occurs weeks after the initial drug-target interaction, and non-specific, as a simple reduction in titre fails to reveal precisely which step of the germinal centre reaction was modulated. To achieve the mechanistic sensitivity required for modern PoM, endpoints must therefore be proximal to the drug's site of action. This necessitates moving beyond blood to the tissues where immune responses are initiated and effector functions enacted, such as the skin, bone marrow, and lymph nodes. However, the value of these approaches is predicated on the ability to reliably sample the target and assay the resulting tissue using robust assays.

1.3.4 Key HIC Agents Considered in this Thesis

To address these limitations, this thesis focuses on refining and standardising models using two specific agents: KLH and MVA-BN.

1.3.4.1 Keyhole Limpet Haemocyanin (KLH)

The xenogenic protein KLH, derived from the Giant Keyhole Limpet (*Megathura crenulata*), is widely considered a model antigen for assessing T-cell dependent antibody responses preclinically and in humans. Its primary advantage lies in its status as a neoantigen; humans are typically immunologically naïve to its epitopes prior to immunisation. This offers a distinct advantage over recall antigens such as Tetanus or Varicella Zoster, where pre-existing immunity varies significantly between individuals based on vaccination history or natural

exposure. Consequently, the use of a neoantigen allows for an end-to-end interrogation of the immune response, from primary sensitisation to memory formation, under strictly controlled conditions.

The KLH challenge protocol typically involves an initial sensitisation phase via intramuscular or subcutaneous injection to induce a primary antibody response.[16] This is followed by a rechallenge phase involving intradermal injection to elicit a Delayed-Type Hypersensitivity response. This pairing allows for the simultaneous assessment of humoral and cell-mediated immunity in accessible tissues. While KLH has been utilised for over 50 years, historical preparations varied significantly in purity, which limited reproducibility. To address this, the studies reported in this thesis utilise a modern GMP-grade subunit KLH formulation (∼400kDa isoforms; Immucothel), which permits the standardisation necessary for robust Proof-of-Mechanism assessment.

1.3.4.2 Modified Vaccinia Ankara (MVA-BN)

MVA-BN is a highly attenuated, non-replicating viral vector that serves as an suitable candidate for a viral challenge model. It combines the potent immunogenicity characteristic of a poxvirus with a strong safety profile, allowing its use even in immunocompromised populations. Similar to KLH, MVA-BN functions effectively as a neoantigen in populations naïve to smallpox vaccination, which includes most individuals born after approximately 1980. However, unlike protein antigens, MVA-BN engages distinct viral sensing pathways and effector mechanisms thereby offering a mechanistically distinct probe of the immune system. Recent clinical interest in MVA-BN due to Mpox outbreaks has led to improved characterisation of its immunological profile, providing a timely opportunity to formalise its use as a standardised Human Immune Challenge tool with rigorous, tissue-based endpoints.

1.4 Quantitative Decision Frameworks

The uncertainty inherent in HIC-derived pharmacodynamic readouts—particularly their imperfect and mechanism-dependent predictive validity—limits the utility of conventional hypothesis-testing frameworks for early-phase decision-making.

While the *Three Pillars* provide the strategic roadmap and HIC models provide the biological substrate, the translation of experimental data into Go/No-go decisions requires a corresponding statistical framework. Historically, the analysis of early-phase clinical trials has relied on traditional frequentist statistical approaches. However, the application of confirmatory statistical standards to exploratory experimental medicine presents specific challenges.

1.4.1 Limitations of Null Hypothesis Significance Testing (NHST)

Conventional frequentist approaches rely on the rejection of a null hypothesis (H_0) to control the Type I error rate (typically $\alpha = 0.05$). While appropriate for confirmatory registration trials where regulatory standards require strict control of false positives, this paradigm prioritises error control over decision value.

In Phase I studies, study sizes are often constrained by ethical and economic factors (typically $n = 12$ to 20 per group). Consequently, strict adherence to standard significance thresholds often results in low statistical power, increasing the risk of Type II error (false negative decisions).

Furthermore, while experienced decision-makers implicitly weigh external evidence (such as the predictive value of the *Three Pillars* data) in a "totality of evidence" approach, the formal statistical analysis of the immediate trial data typically assumes independence from these historical priors. This separation can lead to a disconnect between the statistical

output of a trial and the strategic context of the development program.

1.4.2 Bayesian Decision Theory

Bayesian Decision Theory (BDT) provides a formalised alternative that is aligned with the iterative nature of drug development. Unlike hypothesis testing, which focuses on the statistical significance of an effect, BDT focuses on the consequences of the decision based on the posterior probability of that effect.

By incorporating prior distributions, BDT allows for the mathematical integration of historical success rates (e.g., the 29% vs 0% transition probabilities described by Wu et al.) with the observed trial data. This formalises the integration of prior knowledge that may be otherwise handled qualitatively.

1.4.3 Integrating Predictive Uncertainty and Utility

A central challenge in using HIC for decision-making is the uncertainty of the link between the pharmacodynamic endpoint and the clinical outcome. Unlike validated surrogate endpoints used in later phases, the readouts from HIC (e.g., DTH response) possess uncertain predictive validity regarding clinical efficacy in the target disease.

Current qualitative frameworks address this via expert consensus. However, a quantitative utility-based framework offers distinct advantages by making the risk-benefit assumptions explicit. By modelling the predictive validity of the biomarker as a probability distribution rather than a binary property, and integrating this with the economic consequences of "False Go" versus "False Stop" decisions, BDT allows for the calculation of Net Expected Utility (NU).

This thesis argues that such a framework can reveal optimal decision strategies—such

as the calibration of decision thresholds or sample sizes—that may be counter-intuitive when the problem is viewed solely through the lens of statistical significance or traditional evidence hierarchies. By shifting the objective from maximizing statistical significance to maximizing expected utility, the framework provides a rigorous method for calibrating Go/No-Go decisions in the presence of high biological uncertainty. This thesis therefore contends that Bayesian decision-theoretic approaches are not merely statistically convenient, but economically necessary to extract maximal decision value from HIC-enabled Phase I PoM studies.

1.5 Thesis aims, hypothesis, and objectives

The overarching aim of this thesis was to enhance the feasibility, mechanistic specificity, and decision-making value of HIC models for early-phase PoM assessment in immunomodulatory drug development.

This thesis tests the hypothesis that systematic optimisation of HIC model design—including challenge agent, dose, adjuvant, tissue selection, and assessment timing—enables robust, mechanism-specific PD readouts in healthy volunteers. It further hypothesises that formal integration of these readouts within a Bayesian decision-theoretic framework yields more informative and economically efficient Go/No-Go decisions than conventional Phase I development strategies that either lack mechanism-relevant PD assessment or incorporate such evidence informally. Accordingly, the thesis comprises (i) empirical optimisation and characterisation of HIC agents and tissue-based endpoints, and (ii) a quantitative decision-theoretic evaluation of how such endpoints may be used to support development decisions. This work had five objectives:

Chapter summary

1. To characterise the immune response to a marketed prophylactic vaccine (MVA-BN) in a healthy cohort, to identify predictors of response, and to assess its potential value as a viral HIC agent for fundamental immunology and for PD evaluation in immunomodulatory drug development.

2. To systematically identify, collate, and synthesise HIC studies employing KLH, in order to define design limitations and inform a research agenda for optimisation of KLH HIC paradigms in drug development.

3. To further characterise and optimise the KLH immune challenge model to advance its suitability for early-phase Proof-of-Mechanism studies in healthy volunteers:

   (a) To determine the effect of co-administered adjuvants (Alhydrogel and Montanide ISA-51) on the magnitude and $CD4^+$ T_h1 / T_h2 polarisation of the resultant systemic primary immune response to KLH in healthy volunteers.

   (b) To quantify the effect of KLH rechallenge dose and identify an optimal dose for future PoM studies.

   (c) To characterise the temporal evolution of cutaneous immune responses to KLH rechallenge using multi-parameter flow cytometry, and thereby identify timepoints optimally suited to interrogate mechanistically distinct immune processes.

4. To evaluate, using an experimental medicine approach, the feasibility of

contrast-enhanced ultrasound for identifying primary draining lymph nodes in healthy volunteers to enable tissue-based sampling in HIC and PoM studies.

5. To demonstrate how immune challenge-derived PD readouts may be formally integrated within a Bayesian decision framework, and to examine how utility-calibrated decision rules can optimise Go/No-Go thresholds relative to conventional Phase I decision strategies.

1.5.1 Thesis structure

This thesis is comprised of seven chapters.

Chapter 2 addresses Objective 1 through an observational immunological study of healthy volunteers receiving MVA-BN as mpox prophylaxis.

Chapter 3 addresses Objective 2 via a systematic review of human KLH challenge studies reported in the literature.

Chapter 4 addresses Objectives 3a–c through two KLH HIC studies investigating systemic and cutaneous immune responses following primary immunisation and intradermal rechallenge.

Chapter 5 addresses Objective 4 through an experimental medicine study evaluating the feasibility of contrast-enhanced ultrasound for identifying primary draining lymph nodes in healthy volunteers.

Chapter 6 addresses Objective 5 via a simulation study demonstrating proof-of-concept for Bayesian decision-theoretic integration of KLH HIC-derived PD endpoints in a hypothetical PoM Go/No-Go decision.

Chapter 7 synthesises the main findings of the thesis, discusses their implications for immunomodulatory drug development, and outlines priorities for future research.

Chapter 2

Repurposing of Modified Vaccinia Ankara vaccine: a prospective cohort study and exploration of potential as a novel immune challenge agent

Chapter summary

The replication-incompetent vaccine MVA has several characteristics that support its use as a novel HIC agent including uncommon pre-existing immunity to relevant epitopes in the general population and well-established safety profile. However, to our knowledge it has never been used for this purpose. A recent local implementation of a national prophylactic mpox vaccination programme provided an opportunity to examine the following objectives:

To characterise the immune response to a marketed prophylactic vaccine (MVA-BN) in a healthy cohort, to identify predictors of response, and to assess its potential value as a viral HIC agent for fundamental immunology and for PD evaluation in immunomodulatory drug development.

This chapter is based on a publication in Lancet Microbe (DOI:

10.1016/j.lanmic.2024.101045).[17] Parts of methods and results are substantially reproduced from this work.

2.1 Introduction

Human Immune Challenge (HIC) models are essential tools for dissecting the kinetics of the human immune response and evaluating novel immunomodulatory therapeutics. While live viral challenges (e.g., influenza, malaria) provide a close approximation to natural infection, they require complex biosafety measures and carry distinct safety risks. Conversely, standard vaccines represent a class of HIC that are inherently safe, standardised, and ethically acceptable.

To expand the HIC toolkit, fundamental characterization of vaccine vectors as challenge agents is required. The MVA vector, specifically the Bavarian Nordic formulation (MVA-BN), possesses key attributes that make it suitable as a novel HIC agent. MVA-BN is a highly attenuated, replication-incompetent Orthopoxvirus (OPXV) capable of entering cells and producing viral proteins without generating infectious progeny. This safety profile allows for broader inclusion criteria in studies, including in the setting of pathological or iatrogenic immunosuppression. Furthermore, pre-existing immunity to the MVA vector has until recently been uncommon in the general population following the discontinuation of routine smallpox vaccination coinciding with the eradication of smallpox globally in 1979, reducing the noise of recall responses when establishing baseline immunogenicity.[18]

Despite these favourable characteristics, MVA-BN has not been fully optimised as a standardised immune challenge model. The 2022 global mpox outbreak provided a unique

opportunity to prospectively characterise the immunogenicity of MVA-BN in a healthy cohort. While the clinical impetus for this study was to assess the immunogenicity of dose-sparing intradermal regimens and factors associated with differential immunogenicity, the dataset provides a high-dimensional "systems immunology" view of the agent.

In this chapter, we evaluate MVA-BN not only as a prophylactic measure but as a candidate immune challenge agent. We sought to characterise the distribution and determinants of key serological and cellular outcomes. In so doing we sought to define the operating parameters of MVA-BN for future experimental medicine studies, and identify areas for future optimisation of the model as an agent for HIC.

2.1.1 Aims and Objectives

The study had the following objectives

Chapter Objectives(s)

1. To describe the serological and T cell immune response to MVA-BN vaccination
2. To describe the nature, magnitude, and time course of the primary and secondary immune response to MVA-BN vaccine
3. To describe potential predictors of the immune response to MVA-BN vaccine

2.2 Methods

2.2.1 Study design, participants and sample collection

This was a prospective single-centre observational study of immune responses to MVA-BN vaccine administered to patients attending sexual health MVA-BN vaccination clinics in Oxford, United Kingdom (UK), as part of the public health response to the 2022 outbreak.

The planned recruitment period was from the time of ethical approval through to completion of the initial vaccination programme (approximately 12 months). Clinic attendees were systematically approached by a member of the vaccination team and if interested in the study were invited to participate. People living with HIV (PLWH) were excluded due to limited availability of biosafety category III laboratory facilities for the planned analyses. The recommended MVA-BN dose-regimen was directed by the patient's sexual health physician according to UK Joint Committee on Vaccination and Immunisation (JCVI) guidance. At the time of study enrolment, the JCVI recommended fractionated intradermal vaccination, in the absence of specific contraindications (severe immunosuppression or history of keloid scarring). The intradermal regimen consisted of two doses of 0.1 mL MVA-BN (corresponding to 2×10^7 plaque forming units (PFU), and no less than 1×10^7 PFU per dose), delivered no less than 28 days apart. Individuals who had a history of prior smallpox vaccination were recommended a single MVA-BN 0.1 mL intradermal vaccination. Individuals with contraindications to intradermal administration received standard subcutaneous dosing (0.5 mL MVA-BN, corresponding to 1×10^8 PFU and no less than 0.5×10^8 PFU per dose).

Peripheral blood samples were collected for analysis of serum (serum separator tube, BD), peripheral blood mononuclear cells (PBMCs, lithium heparin tube, BD), and RNA expression (Tempus Blood RNA tubes, ThermoFisher Scientific). Blood samples were processed within 4 hours of collection. PBMCs were isolated by density centrifugation (Lymphoprep, STEMCELL). Serum and plasma samples were stored at $-80\,^{\circ}\mathrm{C}$, and PBMC samples were cryopreserved in 10% dimethylsulfoxide (DMSO) in fetal bovine serum (FBS) in liquid nitrogen until analysis. All laboratory assays were performed in University of Oxford laboratories unless otherwise stated.

Participants had blood sampling at day 0 (prior to vaccination), day 14, and day 28 following their first dose. Those receiving a two-dose regimen had additional blood samples on day 28 and 90 following their second dose ('D2-28' and 'D2-90' respectively). Cryopreserved peripheral blood mononuclear cells and serum were stored prior to analysis at University of Oxford laboratories unless otherwise stated. Participants were also invited to join an 'early response' study cohort, consisting of one additional blood sample on day 1 (i.e. 24 hours) following their first dose—this was offered as an optional study addition due to the short notice from study entry to the day 1 visit. Individuals in the early response cohort were additionally required to be aged 18-45, without medical comorbidities affecting the immune response, and to have no previous smallpox vaccination or history of mpox (suspected or confirmed) based on physician assessment at study enrolment. Baseline demographic characteristics were assessed in all participants, in addition to a medical history focusing on evidence of immunosuppression, prior mpox infection, or prior smallpox vaccination. At each visit, a symptom and exposure questionnaire was performed to assess for evidence of intercurrent mpox infection during follow-up.

The study protocol was approved by the UK NHS Ethics committee (London – Surrey Borders Research REC ref 22/PR/1425). All participants provided written informed consent prior to any study procedures.

2.2.2 Serology

Antibody responses were comprehensively assessed using established assays at UK Health Security Agency (UKHSA) laboratories on a Luminex platform (Diasorin) as described previously.[19, 20] The 'MpoxPlex' assay evaluates IgG antibodies to MPXV proteins B2, A27, B6, M1, A35, E8, A5, H3 & A29 and VACV proteins A27, B5 & A33, with positivity

cut-offs for individual antibody responses, and a scoring system to assign an overall interpretation established in earlier publications.[19, 20] Samples with scores≥4 were considered seropositive. Positive samples with a mean fluorescent intensity (MFI) above the cut-off for MPXV A27 were considered indicative of MPXV infection and/or historical smallpox vaccination (prior to 1972), due to the absence of this protein in the MVA-BN strain but presence in historical smallpox vaccines. Positive samples with MFIs below the cut-off for MPXV A27 were considered indicative of recent smallpox vaccination.

2.2.3 Peptide pools

We used peptide pools as stimuli to detect antigen specific T cell responses following MVA-BN vaccination. Two sets of peptide pools were used: (i) To assess vaccine responses at baseline and expected peak (day 14) T cell response, we used OPXV-specific epitope 'megapools' containing 318 predicted HLA class II-restricted ("OPX-CD4-E") peptides and 238 predicted HLA class I-restricted ("OPX-CD8-E") peptides spanning the vaccinia virus genome (approximately 200 open reading frames). The megapools have been used in recent studies of smallpox vaccination response, demonstrate high sequence homology with corresponding MPXV regions, and are therefore predicted to act as a comprehensive probe of OPXV T cell immunity.[21] (ii) MVA peptide pools containing predicted HLA class II-restricted ("CD4+"; 27 peptides) and HLA class I-restricted ("CD8+"; 36 peptides) peptides spanning the length of MVA CD4+ and CD8+ T cell epitopes from prior MVA-vector vaccine clinical trials were used to longitudinally assess MVA-BN vaccine responses, thus enabling cross-study benchmarking.[22]

A full description of the rationale and methodology for synthesis of the OPXV-specific

megapools is provided in a recent publication from Grifoni et. al.,[21] and a general description of the 'megapool' approach is given in da Silva Antunes et. al.[23] In short, vaccinia virus epitope information from the Immune Epitope Database was used to predict potential HLA class II-restricted and HLA class I-restricted (i.e. CD4+ and CD8+ epitopes respectively) T cell epitopes, which were demonstrated to be highly conserved across OPXV species (including MPXV), thus capable of acting as a broad probe of OPXV T cell immunity following vaccination and infection.[24] Peptides were synthesised as crude material (TC Lab, San Diego, CA), individually resuspended in dimethyl sulfoxide (DMSO) at a concentration of $20\,\mathrm{mg\,mL^{-1}}$. Aliquots of peptides were pooled into 'megapools', designated 'OPX-CD8-E' and 'OPX-CD4-E' for stimulation of CD8+ and CD4+ T cells respectively. The megapools required a sequential lyophilization as previously reported and were resuspended at $1\,\mathrm{mg\,mL^{-1}}$ in DMSO to be used at a final concentration of $1\,\mu\mathrm{g\,mL^{-1}}$.[25]

2.2.4 ELISpot

IFN-γ ELISpot was performed using cryopreserved PBMCs (rested for 2h following thawing), as previously described.[26] For ELISpots, PBMCs (3×10^5/well) were stimulated with the relevant peptide pools or anti-human CD3 mAb (Mabtech AB) as a positive control; media only as negative control. ELISpot plates were incubated for 18-20h at

2.2.5 Flow cytometry and activation induced marker (AIM) assay

Activation induced marker (AIM) assays were used to detect OPXV-specific CD4+ and CD8+ T cell responses in cryopreserved PBMCs, using the OPXV peptide megapools described above. For CD4+ T cells, AIM positivity was defined by upregulation of OX40 (CD134) and 4-1BB (CD137),[27] while for CD8+ T cells, AIM positivity was defined by

upregulation of CD69 and 4-1BB (CD137).[28] AIM responses are reported as the percentage of AIM-positive cells relative to total CD4+ or CD8+ T cells, following background subtraction.

Cryopreserved PBMCs were thawed and incubated at 37 °C with OPXV megapools at a final concentration of $1\,\mu g\,mL^{-1}$ for 20–24 hours in 96-well U-bottom plates containing RPMI 1640 medium supplemented with GlutaMAX™ (Gibco, 61870036) and 10% fetal bovine serum (Sigma Aldrich). Background negative control wells contained PBMCs cultured with 0.2% DMSO (equimolar to the peptide pools). For positive controls, plates were pre-coated with purified anti-human CD3 antibody (OKT3 clone, BioLegend, 317302) at $1\,\mu g\,mL^{-1}$ for 1 hour at 37 °C.

Following stimulation, PBMCs were harvested and resuspended in phosphate-buffered saline containing 2% FBS and $2\,mmol\,L^{-1}$ EDTA (FACS wash buffer). Cells were stained with a surface antibody cocktail prepared in Brilliant Stain Buffer (BD) for 30 minutes at 4 °C in the dark, washed twice, and then fixed and permeabilised using the eBioscience™ Foxp3 / Transcription Factor Staining Buffer Set (Invitrogen, 00-5523-00). Intracellular antibody staining was subsequently performed for 30 minutes at 4 °C in the dark.

All samples were acquired on a Sony ID7000 spectral flow cytometer (software v2.02) and analysed using FlowJo version 10 (FlowJo LLC).

Table 2.1: Antibody list for flow cytometry / AIM assay

Marker	**Clone**	**Fluorescent Tag**	**Species**	**Supplier**	**Catalog no.**
Viability		Zombie NIR		Biolegend	423105
CD14	63D3	BV750	Human	Biolegend	367136
CD56	5.1H11	BV750	Human	Biolegend	362556

CD3	UCHT1	BUV805	Human	BD Biosciences	612894
CD4	SK3	BV510	Human	BD Biosciences	562971
CD8	SK1	Spark Blue 550	Human	Biolegend	344759
CD25	M-A251	PerCP	Human	Biolegend	356131
CD137 (4-1BB)	4B4-1	R718	Human	BD Biosciences	567086
CD69	FN50	BV650	Human	Biolegend	310934
CD40L (CD154)	5C3	PE-Fire 810	Human	Biolegend	310857
CD38	HIT2	BUV661	Human	BD Biosciences	612969
OX40	ACT35	BV605	Human	Biolegend	350028
PD-1	MIH4	PE-CF594	Human	BD Biosciences	566846
PD-L1	MIH5	BUV737	Human	BD Biosciences	741881
CD127	A019D5	PE-Fire 640	Human	Biolegend	351363
HLA-DR	L243	BUV563	Human	BD Biosciences	753686
CD27	O323	BUV496	Human	BD Biosciences	751678
CD28	CD28.2	BB515	Human	BD Biosciences	564493
CD45RA	HI100	Spark YG 581	Human	Biolegend	304173
CD19	HIB19	cFluor B532	Human	Cytek Biosciences	R7-20072
IgD	IA6-2	BV480	Human	BD Biosciences	566187
IgM	SA-DA4	BV786	Human	BD Biosciences	740998
CD138	MI15	BUV615	Human	BD Biosciences	751148
IgG	G18-145	PE-Cy7	Human	BD Biosciences	561298
CD71	L01.1	RB780	Human	BD Biosciences	755371
T-bet	4B10	Alexa Fluor 647	Human	Biolegend	644803
GATA3	L50-823	BUV395	Human	BD Biosciences	565448
RORyt	Q21-559	PE	Human	BD Biosciences	563081
FOXP3+	206D	BV421	Human	Biolegend	320123
IRF4	3E4	PerCP-eFluor710	Human	Invitrogen	46-9858-82

Ki-67	Ki-67	BV711	Human	Biolegend	350515

2.2.6 Statistical analysis

Statistical analysis was performed using R (v4.4.0) in RStudio (2024.04.1+748 for macOS) unless stated otherwise. A p-value of 0.05 was used to assess statistical significance. The cohort and resultant immune responses were described using descriptive and exploratory comparative statistics, and analyses considered potential confounding demographic and immunological characteristics (e.g. CMV serostatus,[29] recent vaccination or infection). Non-parametric techniques were used for univariable hypothesis tests (Wilcoxon rank sum test). For analysis of serological responses (log10 transformed MFI value for each antigen target), Generalised Linear Mixed Modelling was used (lme4 v1.1.35.3) to account for repeated measures. For T cell responses, the primary metric of interest was the difference (delta) in T cell frequency ($SFC/10^6$ PBMC for ELISpot, and % AIM+ positive cells for AIM assay) between day 0 and day 14. We also summarised an individual 'positive' response to vaccination based on an empirical threshold of >2 fold increase in antigen specific T cell frequency between these timepoints.

2.2.7 Generalised Linear Mixed Modelling of serological responses

2.2.7.1 Serological responses accounting for effect of prior vaccination and other covariates

Antibody responses were modelled using Generalised Linear Mixed Modelling to account for repeated within-participant assessment of antibody responses over time.

First, we assessed the effects of potential covariates (previous vaccination, age, BMI, and CMV serostatus) on antibody response. These covariates were assumed to affect all antigen

targets in the UKHSA MpoxPlex panel equally. We therefore modelled the effect of these covariates on the mean log(MFI) response, restricted to only those antibodies which were observed to rise following vaccination, based on a positive linear trend between time and antibody response ($p < 0.05$). These antibodies were termed 'dynamic' to distinguish them from antibody responses to targets in the UKHSA MPoxPlex panel which did not appear to rise following vaccination. We first fit a simple linear regression of mean antibody response according to time and previous vaccination status. We next assessed random intercepts and slopes according to participant ID. Finally, we assessed the effect of age, BMI, and CMV serostatus on model fit. At each step the Akaike information criterion was used to assess for model fit, where a decrease in AIC of 3.84 or greater was used to define improvement. Any significant covariates retained in these steps were selected as covariates for the subsequent modelling of individual antibody responses (described below).

Next, responses to each antigen in the UKHSA MpoxPlex panel were fit separately using an iterative approach beginning with a simple linear model, followed by assessment of random intercepts and slopes according to participant ID. At each step the Akaike information criterion was used to assess for model fit, where a decrease in AIC of 3.84 or greater was used to define improvement.

For these mixed effects models uncertainty intervals (equal-tailed) and p-values (two-tailed) were computed using a Wald t-distribution approximation using the R package parameters (0.22.1). Uncertainty intervals for random effect variances were computed using a Wald z-distribution approximation using the R package parameters (0.22.1).

2.2.7.2 Serological response to second MVA-BN vaccination stratified by day 28 response

To assess the serological response to a second MVA-BN in individuals according to day 28 serostatus, we used an equivalent generalised linear mixed modelling approach to that described in 2.2.7.1, this time limited to dynamic antibody responses from day 28 onward.

2.3 Results

2.3.1 Cohort description

One hundred and ten individuals presented for MVA-BN vaccination during the study recruitment period (1 December 2022 – 3 May 2023), of whom 35 were recruited (figure 2.1). One participant withdrew consent prior to any sampling and was excluded from follow-up. Eighteen of 34 participants (53%) enrolled in the early response cohort (day 1 blood sample). The median participant age was 32 (25th-75th percentile 27 - 42.75) and 33 (97%) were male. Twenty-nine of 34 participants (85.3%) received a two-dose fractionated (0.1 mL) intradermal regimen, with the second dose administered 28 days following the first dose (range 26-40 days). One participant received a two-dose standard (0.5 mL) subcutaneous regimen due to contraindication to intradermal injection and was included in subsequent analyses—a sensitivity analysis excluding these data demonstrated no substantive effects on study conclusions. Four participants (two with visible vaccine scars) had a history of historical smallpox vaccination (estimated year of vaccination 1962-1976) and thus received a single fractionated intradermal dose. These individuals were older than those without a history of vaccination (median age 57 (range 51-72) vs 30 (23-49) years). Nineteen of 34 participants (59.4%) were cytomegalovirus (CMV) seropositive.

No individuals reported a history consistent with symptomatic mpox infection prior to study entry, or during follow-up. Participants receiving two doses (n=30) were followed for a minimum of 28 days following dose 1, with 28 (93%) followed through to D2-90. Two individuals missed the day 14 sampling timepoint. All participants receiving single dose regimens (n=4) were followed to D1-90.

Characteristic	**Seronegative**	**Seropositive**
n	16	14
Age (median [25th–75th percentile])	29 [27, 41]	32 [26, 34]
BMI (median [25th–75th percentile])	25.45 [21.4, 27.3]	23.95 [22.5, 26.2]
CMV seropositive (%)	6 (42.9)	10 (71.4)
Current smoker (%)	0 (0.0)	3 (21.4)
Vaccination within 30 days of study entry (%)	2 (14.3)	2 (15.4)
COVID-19 infection within 30 days of study entry (%)	14 (100.0)	13 (100.0)
Sexually transmitted infection treated within 30 days of study entry (%)	1 (7.1)	1 (7.7)

Table 2.2: Cohort characteristics stratified by day 28 serostatus. No statistically significant differences were observed in these characteristics according to day 28 serostatus (Chi-squared test for categorical variables, Wilcoxon signed rank test for continuous variables).

2.3.2 Serological outcomes

To determine the sensitivity of serology as a primary endpoint for an MVA-BN challenge model, we assessed the dynamic range of antibody responses. All participants without a history of MVA-BN or smallpox vaccination were seronegative at baseline (MpoxPlex definition). Of the 12 targets used in the assay, seven (B2, H3, E8, M1, A35, V.B5, and V.A33) demonstrated a statistically significant rise between day 0 and day 28 (p-value for linear association between time and antibody response $\leq$0.002) (figure 2.2). All seven antigens demonstrated evidence for a boosting effect after dose two. One individual who seroconverted by D2-90 also seroconverted to MPXV A27, a target not present in MVA-

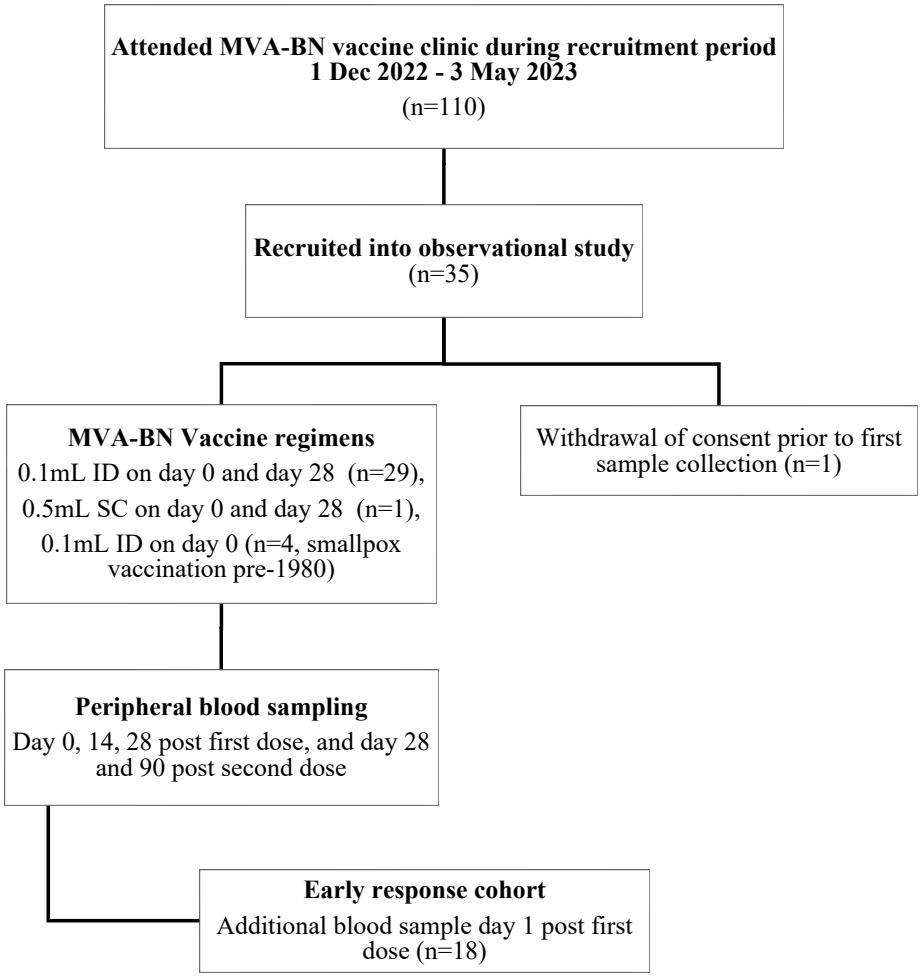

Figure 2.1: Flow diagram for the IMOVA study cohort. Overview of the number of participants in the study, timepoints of peripheral blood sampling and vaccine regimens received. Image reproduced from [17].

BN, potentially indicative of mpox infection, although this participant had no evidence of symptomatic mpox disease or mpox contact. At day 28 post first vaccination, 47% of participants were seropositive, increasing to 89% by D2-90 (figure 2.2). Of those who were seronegative at day 28 (post first dose), 80% had seroconverted by D2-90. The breadth of responses at this timepoint was lower in those individuals who were seronegative at day 28, with a median MPoxPlex score of 5 (range 2-6) versus 7 (4-7) positive antibody responses,

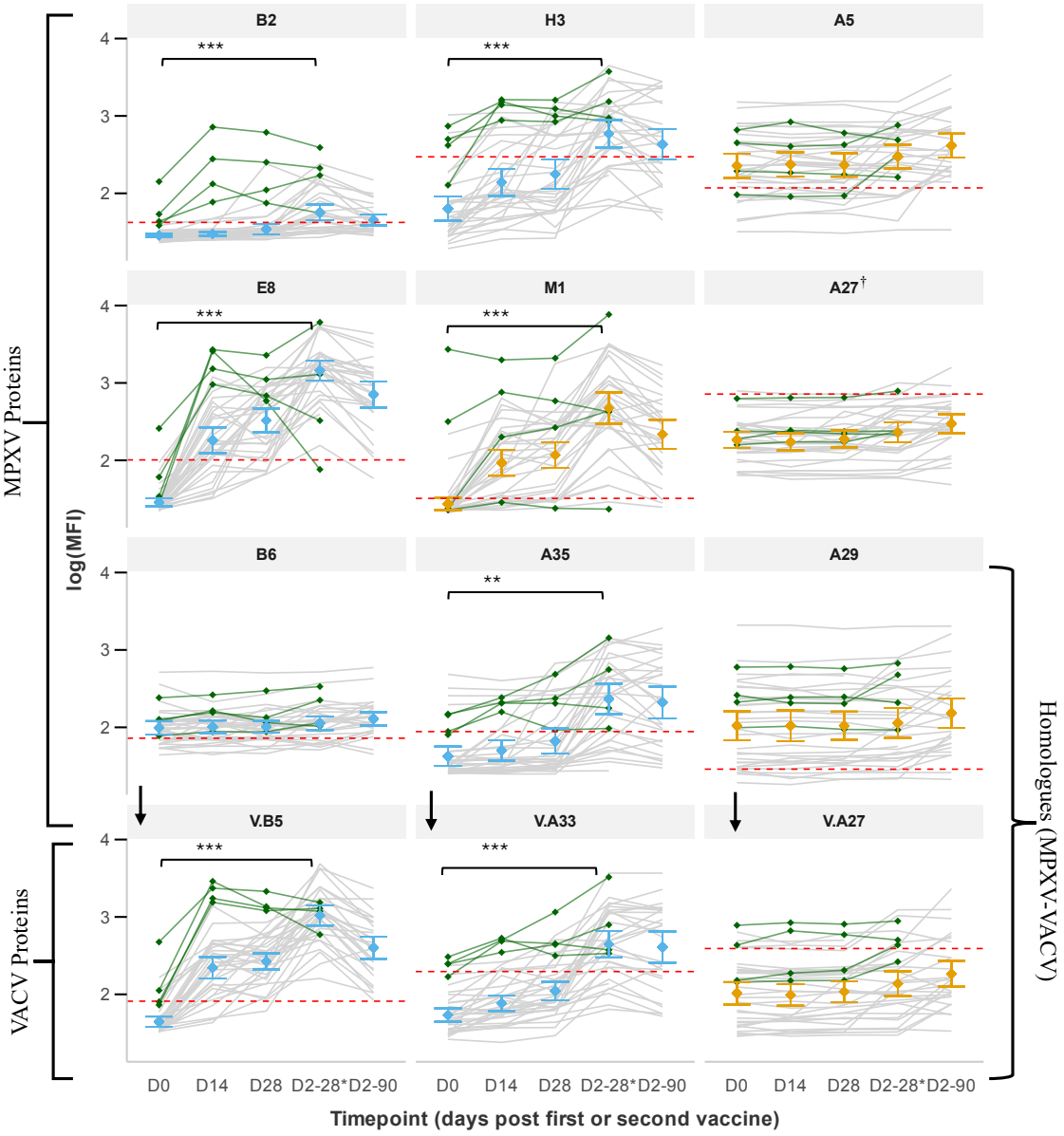

Figure 2.2: Antibody responses to MVA-BN vaccination. Time course of serological responses to each antigen target in the UKHSA MpoxPlex assay at each timepoint. Mean response (log10 transformed median fluorescence intensity, MFI) and 95% confidence intervals are displayed (blue: included in MpoxPlex serological score, orange: not included). 'D2-28' and 'D2-90' indicate the post second dose day 28 and 90 day timepoints respectively. Grey lines represent individuals (n=30) without prior vaccination. Individuals with historical smallpox vaccination (n=4) are shown in green – for these individuals the timepoint 'D2-28' corresponds to 90 days post vaccine. Connecting arrows refer to MPXV-VACV homologous proteins. †MPXV A27L is not a MVA-BN antigen but indicative of mpox infection or historical smallpox vaccination. *$p < 0.05$, **$p < 0.01$, *** $p < 0.001$ for linear effect of time on antibody response for timepoints between D0 and D1-28 (generalised linear mixed effects model). Image reproduced from [17].

and the response to six of the seven dynamic antigens was numerically lower ($p < 0.05$, figure 2.3).

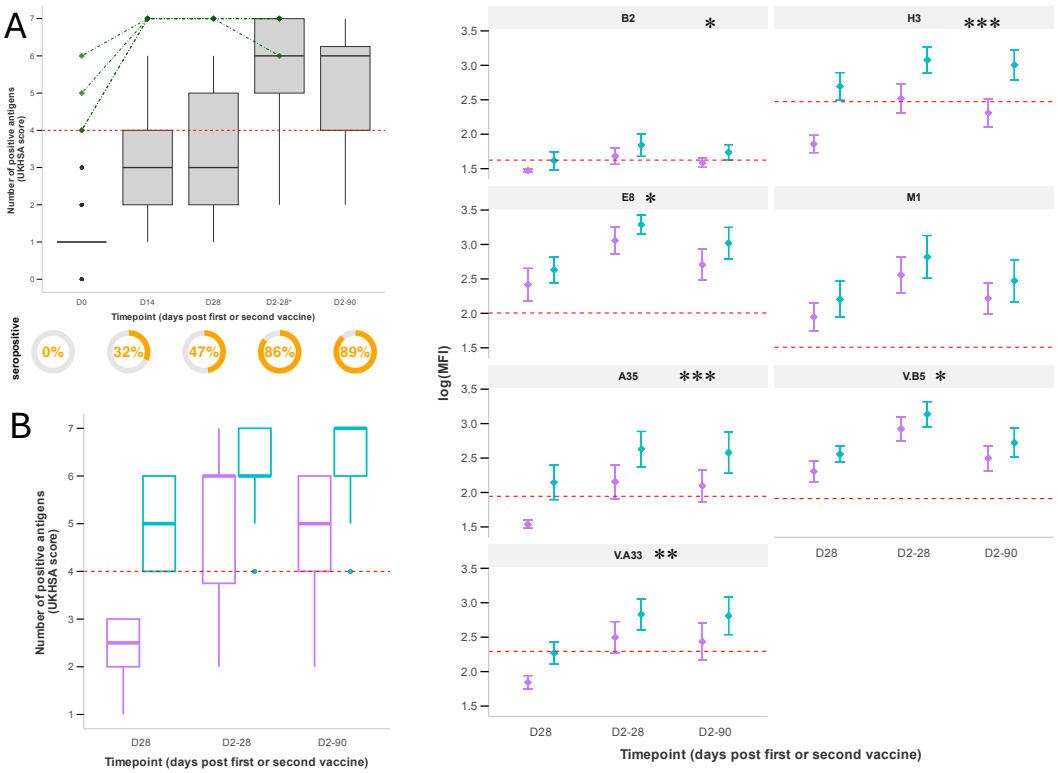

Figure 2.3: Antibody responses to MVA-BN vaccination in participants without previous MVA/smallpox vaccination. A) Boxplots showing time course of UKHSA MpoxPlex serological score (range 0-7, prespecified positive response cut-off = 4) and % seropositivity for individuals (n=30) with no history of MVA/smallpox vaccination. Serological scores for individuals with a history of smallpox vaccination (n=4) are overlaid in green. B) Boxplots showing time course of MpoxPlex serological score from day 28 onwards for individuals with no history of MVA/smallpox vaccination, categorised by day 28 serostatus (purple: seronegative, blue: seropositive) – breadth of responses (number of positive antigens) for day 28 seronegative individuals remains lower than seropositive individuals despite second vaccination administered on / around day 28. C) Responses to antibodies demonstrating dynamic change (significant linear association between time and antibody response as per analysis depicted in figure 2.2), stratified by day 28 serostatus (purple: seronegative, blue: seropositive) – decreased antibody response as defined by serological status at day 28 persists following second vaccination. *$p < 0.05$, **$p < 0.01$, *** $p < 0.001$ for association between day 28 serostatus and log10(MFI) accounting for linear effect of time from day 28 onwards. Image reproduced from [17].

2.3.3 T cell responses

We next assessed peak (day 14) T cell responses via IFN-γ ELISpot and AIM assay (CD4+ and CD8+ OPXV megapools, figure 2.5 and 2.4). At day 14, 88% of participants demonstrated a positive CD4+ T cell response, with a median frequency of 124 (range 1.33-1057) $SFC/10^6$ PBMC, while 92% demonstrated a positive CD8+ T cell response with a median frequency of 245 (0-897) $SFC/10^6$ PBMC. For CD4+ T cells, the median AIM increase was 0.01% (range -0.004 – 0.14, $p < 0.0001$), with 76% demonstrating at least 2-fold increase in the percentage of AIM positive cells. For CD8+ T cells the median delta was 0.13% (range -0.01 – 1.35%, $p < 0.0001$), with 93% of participants demonstrating at least 2-fold increase in percentage of AIM positive cells. Responses to the MVA peptide pools were qualitatively similar to the OPXV megapool responses (figure 2.6). In summary, the induction of a robust CD8+ T cell response highlights the utility of MVA-BN as a model for interrogating MHC Class I-restricted pathways.

Characteristic	**Seronegative (n=16)**	**Seropositive (n=14)**
MVA CD4, $SFC/10^6$ PBMC (median [range])	21.33 [0.00, 612.00]	38.67 [12.00, 146.00]
MVA CD8, $SFC/10^6$ PBMC (median [range])	54.00 [1.33, 281.33]	168.00 [0.00, 1021.33]
OPXV CD4, $SFC/10^6$ PBMC (median [range])	189.33 [38.67, 1057.33]	102.00 [1.33, 320.00]
OPXV CD8, $SFC/10^6$ PBMC (median [range])	216.00 [34.67, 686.67]	257.67 [0.00, 897.33]

Table 2.3: Day 14 T cell response (IFN-γ ELISpot) stratified by day 28 serostatus. MVA = standard MVA peptide pools, OPXV = OPXV mega-pools. No statistically significant differences were observed in any ELISpot response according to day 28 serostatus (Wilcoxon signed rank test).

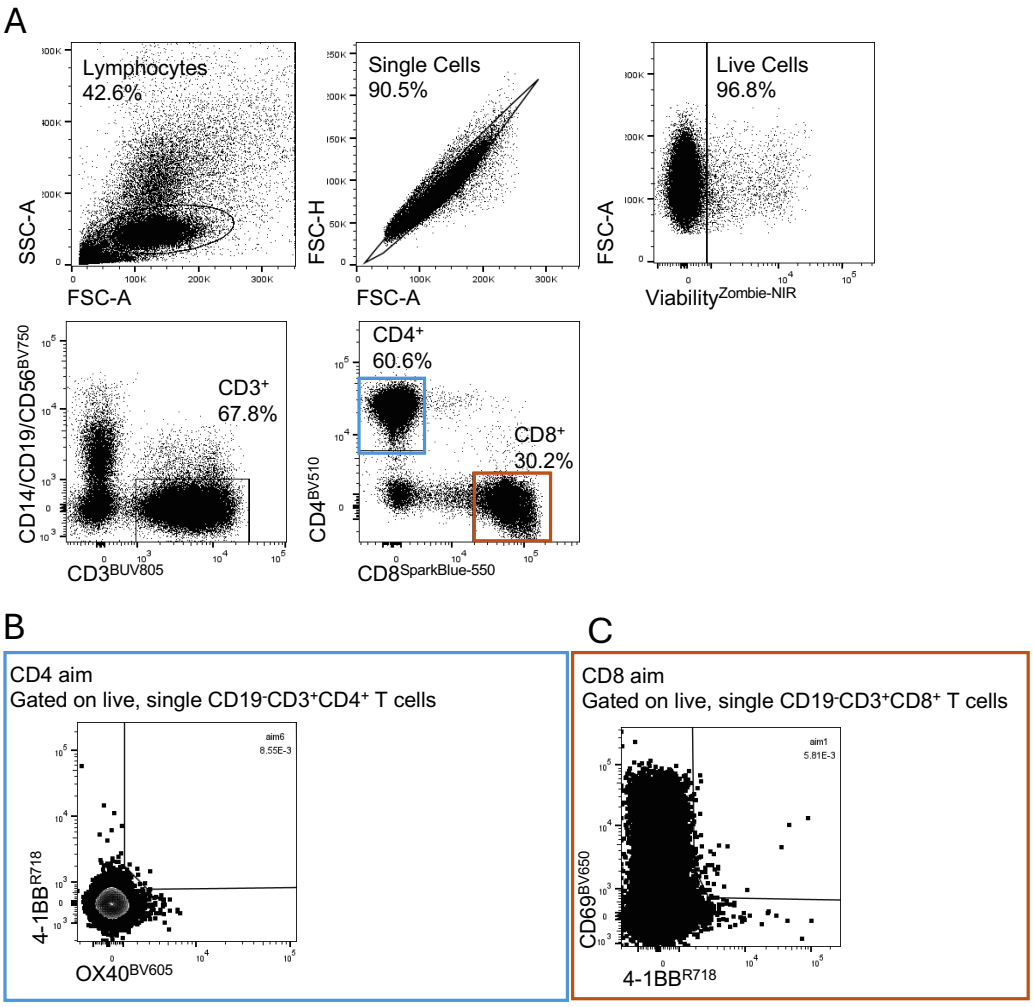

Figure 2.4: Flow cytometry gating strategy for activation induced marker analysis . A. Representative dot plots showing the gating strategy of PBMCs challenged at day 14 post-vaccination with OPXV megapools specific for CD4+ T cells. Representative dot plots of B) CD4, and C) CD8 T cell populations challenged at day 14 post-vaccination with OPXV megapools specific for CD4+ or CD8+ T cells respectively. AIM+ gates were set based on unstimulated samples. Analysis performed by Dr Roel deMaeyer. Image reproduced from [17].

2.3.4 Comparison of T cell responses elicited using OPXV megapools versus standard MVA peptide pools

When compared with the MVA peptide pools, peak (day 14) responses to the megapools were of greater magnitude (CD4+ median 30 [range 0-612] versus 124 [1.33-1057] $SFC/10^6$ PBMC, CD8+ 84.7 [0-1021] vs 245 [0-897] $SFC/10^6$ PBMC. Resultantly, the proportion

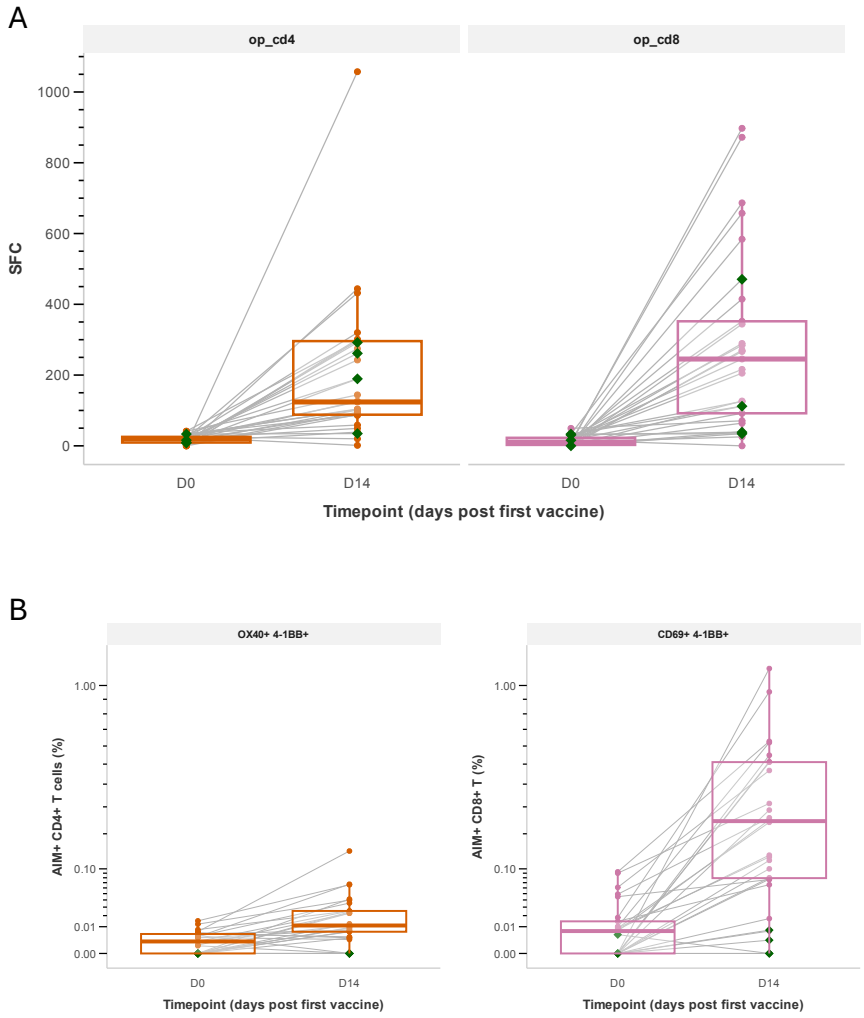

Figure 2.5: T cell responses to MVA-BN vaccination. A) IFN-γ ELISpot T cell responses to stimulation with OPXV peptide megapools (left CD4+ , right CD8+) at day 0 and day 14 post MVA-BN vaccination. Individuals with historical smallpox vaccination are shown in green. Results expressed as spot forming cells (SFC) per 10^6 PBMC. B) Activation induced marker (AIM) positive T cell responses (left CD4+ and right CD8+) at day 0 and day 14 post MVA-BN vaccination. Individuals with historical smallpox vaccination are shown in green. Results expressed as % of AIM positive cells of total CD4+ and CD8+ cells respectively. Image reproduced from [17].

of individuals demonstrating a > 2 fold increase in IFN-γ ELISpot response at day 14 post was higher for the OPXV megapools than for the standard MVA pools (CD4+ 88% vs 56.5%, p = 0.03, $CD8^+$ 92% vs 81.8%, p = 0.5), demonstrating the enhanced sensitivity of the megapools for elicitation of antigen specific T cell response to MVA vaccination.

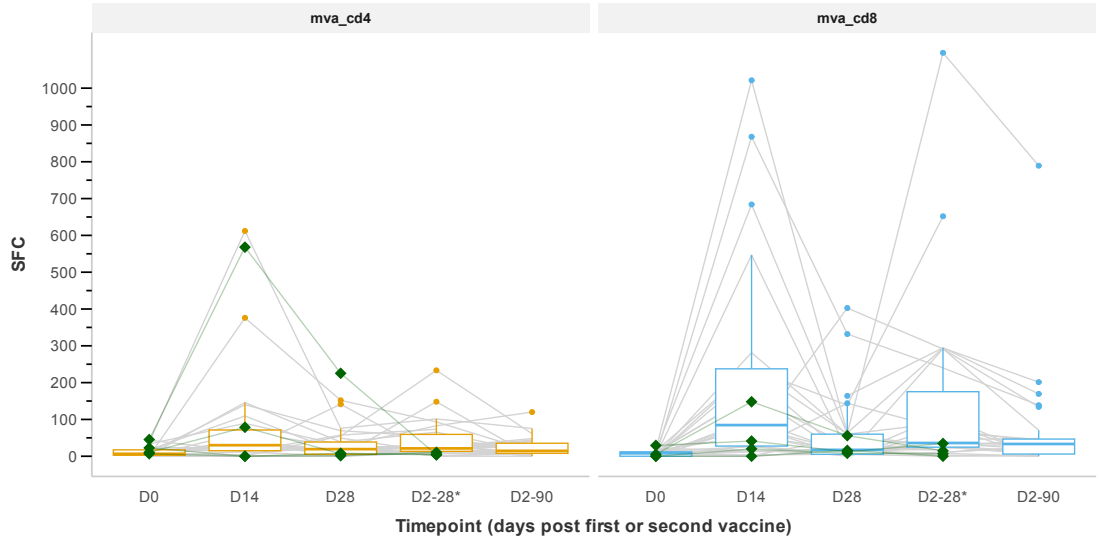

Figure 2.6: Time course of ELISpot responses to MVA peptide pools (orange: CD4+, blue: CD8+) in participants without history of smallpox vaccination (n=30). Individuals with historical smallpox vaccination (n=4) are shown in green – for these individuals the timepoint marked D2-28 corresponds to 90 days post vaccine. Image reproduced from [17].

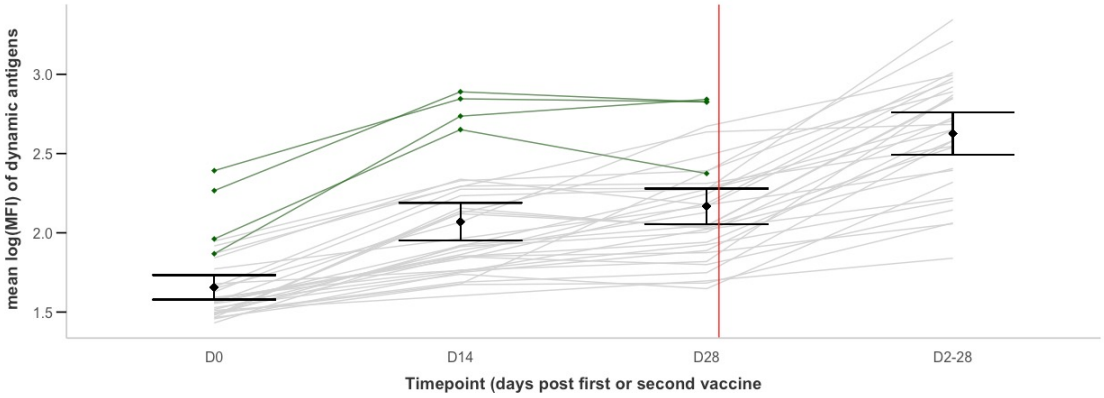

Figure 2.7: Mean antibody response over time, restricted to the seven antigen targets demonstrating a statistically significant rise following vaccination (B2, H3, E8, M1, A35, V.B5, and V.A33; termed dynamic antigens). For each participant, the mean $\log_{10}$(MFI) was calculated across these seven targets at each timepoint. Black dots represent the group mean of individual participant means; error bars represent ± 1 standard deviation. Individuals with a history of smallpox vaccination are shown in green; those without prior vaccination are shown in grey. Image reproduced from [17].

In contrast to the MVA peptide pools, a less pronounced skew was observed on ELISpot performed using the megapools (median CD8+ vs CD4+ response of 245 vs 124 $SFC/10^6$ PBMC, $p = 0.25$) with a CD8+:CD4+ ratio > 1 observed in 56% of participants, suggesting

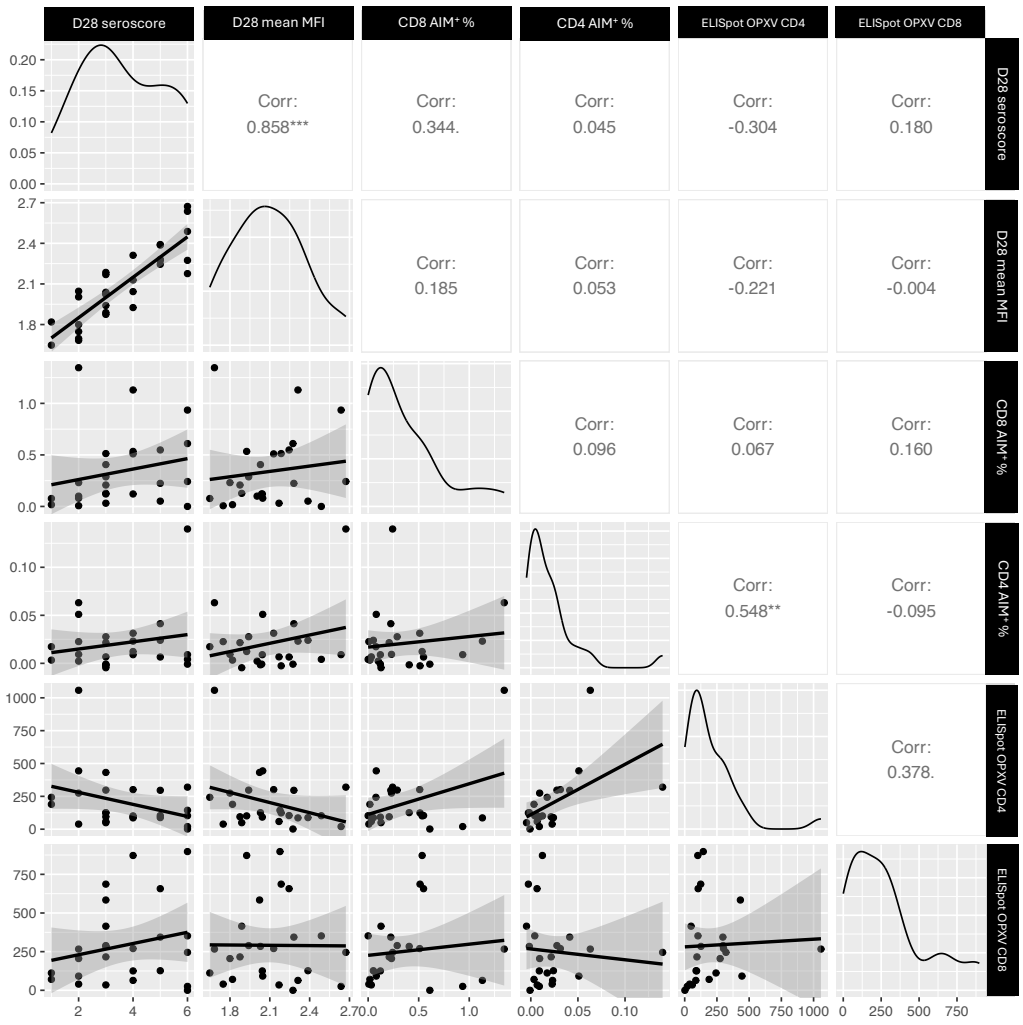

Figure 2.8: Correlations between Day 28 MpoxPlex serology score, serology mean log MFI (dynamic antigens only) and selected day 14 (peak) T cell response metrics. Diagonal plots: density distribution of metric on corresponding x/y axis. Lower off-diagonal plots: scatterplot of individual responses according to corresponding x and y axis metrics. Upper off-diagonal: Spearman correlations between corresponding x and y axis metrics. **indicates p< 0.01, ***< 0.001, all other correlations not statistically significant. ELISpot responses are expressed as $SFC/10^6$ PBMC, AIM expressed as % of CD4+ or CD8+ T cells. Image reproduced from [17].

more balanced CD8+ vs CD4+ responses when T cell immunity is assessed using the OPXV megapools.

2.3.5 Baseline characteristics associated with immunogenicity of MVA-BN and associations between immune metrics

Of the 30 prior smallpox vaccine naïve individuals in the cohort, there was no statistically significant association between age, CMV serostatus, history of sexually transmitted infection or routine vaccination within 30 days of study entry and either serological response or T cell response. Pairwise correlations between Day 28 serological score, mean log MFI across dynamic antigens, and peak T cell responses (ELISpot and AIM) are shown in Figure 2.8. Serological and T cell outcomes were weakly and non-significantly correlated.

2.3.6 Immune responses in individuals with a history of smallpox vaccination

All four participants with a history of smallpox vaccination were seropositive at baseline and their responses appeared broader and of greater magnitude than individuals without previous vaccination (figures 2.2 and 2.7). In contrast, T cell responses did not differ relative to the vaccine naïve individuals (figure 2.6).

All four participants with a history of smallpox vaccination were seropositive at baseline, (median [range] positive antigens 4.5 [4-6]). Serological responses to six antigens (B2, H3, E8, A35, V.B5, V.A33, $p < 0.02$) were significantly higher than the unvaccinated individuals at baseline. The number of positive antigens increased following vaccination in all cases, with all four individuals seropositive to all seven antigens in the UKHSA score by day 28. These individuals mounted higher antibody responses than those without prior vaccination to six of 12 antigens, which largely overlapped with the antigens significantly elevated at baseline (B2, H3, E8, V.B5, V.A33, V.A27, $p < 0.05$). Together this suggests, broader, more rapid, and quantitatively higher antibody responses following a single (booster) vaccination

in those with historical smallpox vaccination compared to the vaccine naïve group. In contrast, T cell responses did not differ systematically in the previously vaccinated group relative to the rest of the cohort (figures 2.2 and 2.5): there was no detectable CD4+ AIM response in any of these individuals, and minimal CD8+ response (median delta 0.001%, range -0.005 – 0.008%). The immunological characteristics of these individuals is consistent with previous studies which show a striking persistence of antibody response decades following prior smallpox vaccination.

2.3.7 Longitudinal ELISpot responses

We next investigated the temporal dynamics of T cell responses to the standard MVA peptide pool at all timepoints using IFN-γ ELISpot. Responses to both CD4+ and CD8+ MVA peptides significantly increased between day 0 and the peak response observed at day 14 (CD4+: median 6.7 $SFC/10^6$ PBMC [range 0-44] day 0 versus 30 [0-612] $SFC/10^6$ PBMC at day 14, $p = 0.006$; CD8+: 8 [0-25.3] $SFC/10^6$ PBMC at day 0 versus 84.7 [0-1021.3] $SFC/10^6$ PBMC at day 14, $p < 0.001$). A second peak response was observed 28 days following the second vaccine (CD4+: 22 [0-233.3] $SFC/10^6$ PBMC, CD8+: 66.7 [2.7-1096] $SFC/10^6$ PBMC, both $p < 0.001$ relative to day 0). CD8+ responses were significantly higher than CD4+ ($p = 0.02$), with a CD8+:CD4+ ratio > 1 at day 14 in 71% of participants.

2.3.8 Mean antibody responses accounting for effect of prior vaccination and other covariates

There was a strong association between mean antibody response and both time and previous vaccination ($p<0.0001$). There was no association between mean antibody response and age,

BMI, or CMV serostatus based on AIC ($p > 0.05$ for these variables).

The final `lme4` model specification was as follows:

$$\text{log\_value} \sim \text{tp} + \text{prevvax} + \text{vaxperiod} + (1|\text{ID})$$

Where:

- **log_value**: Log_{10} transformed antibody response (MFI value).
- **tp**: Time period (months since first vaccination).
- **prevvax**: Whether the participant received prior smallpox vaccination (0 = no, 1 = yes).
- **vaxperiod**: Indicator variable corresponding to the time point, to account for the potential effect of the second vaccine (0 = day 0–28, 1 = day 28 post second vaccine onwards). **Note:** This parameter must be interpreted relative to the linear trend (parameter value for **tp**). If non-significant, this suggests a linear trend for the effect of time (**tp**) between day 0 and day 2–28 is appropriate. If significant, this suggests non-linearity of the response. For example, if **vaxperiod** is negative and its absolute value is greater than **tp** (and statistically significant), this suggests a waning of antibody responses at day 28 post second vaccine.
- **ID**: Participant ID.

Table 2.4: Mixed-effects model parameter estimates for $\log_{10}$-transformed antibody responses (MFI) to individual antigens in the UKHSA MpoxPlex assay. Each antigen was modelled separately. Fixed effect coefficients represent the estimated change in $\log_{10}$(MFI) per unit change in the predictor: tp = time in months since first vaccination; prevvax = prior smallpox vaccination (1 = yes); vaxperiod = indicator for the post-second-dose period (1 = day 28 post second dose onwards). Random effects represent between-participant variability in baseline antibody level (SD Intercept) and residual within-participant variability (SD Residual). Uncertainty intervals and p-values were computed using a Wald t-distribution approximation.

Parameter	Coefficient	SE	95% CI	t(123)	p
A5 - Fixed Effects					
(Intercept)	2.36	0.08	[2.21, 2.51]	31.11	< .001
tp	5.95e-03	0.03	[-0.05, 0.06]	0.21	0.835
prevvax	0.06	0.22	[-0.37, 0.49]	0.28	0.783
vaxperiod	0.09	0.05	[-0.01, 0.19]	1.85	0.067
A5 - Random Effects					
SD (Intercept: ID)	0.40	0.05	[0.31, 0.52]		
SD (Residual)	0.12	8.60e-03	[0.10, 0.14]		
A27 - Fixed Effects					
(Intercept)	2.26	0.06	[2.15, 2.37]	40.51	< .001
tp	0.01	0.02	[-0.03, 0.06]	0.54	0.593
prevvax	0.17	0.16	[-0.14, 0.49]	1.07	0.287
vaxperiod	0.08	0.04	[0.00, 0.15]	2.03	0.045
A27 - Random Effects					
SD (Intercept: ID)	0.29	0.04	[0.23, 0.38]		
SD (Residual)	0.09	6.66e-03	[0.08, 0.10]		
A29 - Fixed Effects					

Continued on next page

Parameter	Coefficient	SE	95% CI	t(123)	p
(Intercept)	2.02	0.09	[1.84, 2.20]	21.79	< .001
tp	-3.20e-03	0.02	[-0.04, 0.03]	-0.17	0.868
prevvax	0.35	0.27	[-0.18, 0.88]	1.31	0.194
vaxperiod	0.06	0.03	[0.00, 0.13]	1.90	0.060
A29 - Random Effects					
SD (Intercept: ID)	0.50	0.06	[0.39, 0.64]		
SD (Residual)	0.08	5.81e-03	[0.07, 0.09]		
A35 - Fixed Effects					
(Intercept)	1.61	0.07	[1.47, 1.75]	22.47	< .001
tp	0.21	0.07	[0.07, 0.34]	3.09	0.002
prevvax	0.51	0.18	[0.15, 0.88]	2.78	0.006
vaxperiod	0.34	0.12	[0.11, 0.57]	2.89	0.005
A35 - Random Effects					
SD (Intercept: ID)	0.31	0.05	[0.23, 0.41]		
SD (Residual)	0.28	0.02	[0.24, 0.32]		
B2 - Fixed Effects					
(Intercept)	1.43	0.04	[1.36, 1.50]	40.30	< .001
tp	0.13	0.04	[0.05, 0.21]	3.36	0.001
prevvax	0.63	0.09	[0.46, 0.80]	7.36	< .001
vaxperiod	0.06	0.07	[-0.07, 0.20]	0.95	0.344
B2 - Random Effects					

Continued on next page

Parameter	Coefficient	SE	95% CI	t(123)	p
SD (Intercept: ID)	0.13	0.02	[0.09, 0.19]		
SD (Residual)	0.16	0.01	[0.14, 0.18]		
B6 - Fixed Effects					
(Intercept)	1.99	0.04	[1.91, 2.08]	48.75	< .001
tp	0.02	0.02	[-0.02, 0.06]	0.76	0.448
prevvax	0.15	0.12	[-0.08, 0.38]	1.30	0.196
vaxperiod	0.03	0.04	[-0.04, 0.10]	0.84	0.401
B6 - Random Effects					
SD (Intercept: ID)	0.21	0.03	[0.16, 0.27]		
SD (Residual)	0.08	6.11e-03	[0.07, 0.10]		
E8 - Fixed Effects					
(Intercept)	1.53	0.06	[1.41, 1.66]	23.70	< .001
tp	1.08	0.09	[0.90, 1.26]	11.84	< .001
prevvax	0.61	0.13	[0.34, 0.87]	4.57	< .001
vaxperiod	-0.52	0.16	[-0.84, -0.21]	-3.31	0.001
E8 - Random Effects					
SD (Intercept: ID)	0.13	0.06	[0.05, 0.30]		
SD (Residual)	0.37	0.03	[0.32, 0.43]		
H3 - Fixed Effects					
(Intercept)	1.83	0.08	[1.67, 1.99]	22.77	< .001

Continued on next page

Parameter	Coefficient	SE	95% CI	t(123)	p
tp	0.45	0.08	[0.29, 0.61]	5.55	< .001
prevvax	0.86	0.20	[0.46, 1.26]	4.24	< .001
vaxperiod	0.04	0.14	[-0.24, 0.32]	0.31	0.758
H3 - Random Effects					
SD (Intercept: ID)	0.33	0.05	[0.24, 0.45]		
SD (Residual)	0.33	0.02	[0.29, 0.39]		
M1 - Fixed Effects					
(Intercept)	1.52	0.09	[1.35, 1.69]	17.50	< .001
tp	0.59	0.08	[0.44, 0.75]	7.48	< .001
prevvax	0.56	0.23	[0.11, 1.00]	2.47	0.015
vaxperiod	-0.04	0.14	[-0.32, 0.23]	-0.31	0.754
M1 - Random Effects					
SD (Intercept: ID)	0.38	0.06	[0.28, 0.51]		
SD (Residual)	0.33	0.02	[0.28, 0.38]		
V.A27 - Fixed Effects					
(Intercept)	2.00	0.07	[1.86, 2.15]	27.58	< .001
tp	0.03	0.03	[-0.03, 0.09]	0.85	0.397
prevvax	0.50	0.21	[0.09, 0.91]	2.42	0.017
vaxperiod	0.10	0.05	[0.00, 0.21]	1.98	0.050
V.A27 - Random Effects					
SD (Intercept: ID)	0.38	0.05	[0.30, 0.49]		

Continued on next page

Parameter	Coefficient	SE	95% CI	t(123)	p
SD (Residual)	0.13	9.21e-03	[0.11, 0.14]		
V.A33 - Fixed Effects					
(Intercept)	1.73	0.06	[1.62, 1.84]	31.07	< .001
tp	0.31	0.06	[0.19, 0.44]	5.06	< .001
prevvax	0.70	0.13	[0.43, 0.96]	5.19	< .001
vaxperiod	0.29	0.11	[0.07, 0.50]	2.64	0.009
V.A33 - Random Effects					
SD (Intercept: ID)	0.20	0.04	[0.14, 0.29]		
SD (Residual)	0.26	0.02	[0.22, 0.30]		
V.B5 - Fixed Effects					
(Intercept)	1.73	0.06	[1.62, 1.84]	30.60	< .001
tp	0.81	0.08	[0.66, 0.96]	10.68	< .001
prevvax	0.74	0.12	[0.50, 0.98]	6.06	< .001
vaxperiod	-0.33	0.13	[-0.59, -0.07]	-2.49	0.014
V.B5 - Random Effects					
SD (Intercept: ID)	0.14	0.04	[0.08, 0.25]		
SD (Residual)	0.31	0.02	[0.27, 0.36]		

2.3.9 Antibody trends stratified by day 28 serostatus (dynamic antigens only)

The final `lme4` model specification was as follows:

$$\text{log\_value} \sim \text{tp} + \text{prevvax} + \text{vaxperiod} + \text{serostatus} + (1|\text{ID})$$

Where:

- **serostatus**: Day 28 serostatus according to UKHSA definition, 0 = seronegative, 1 = seropositive.
- **log_value**: Log_{10} transformed antibody response (MFI value).
- **tp**: Time period (months since second vaccination).
- **prevvax**: Whether the participant received prior smallpox vaccination (0 = no, 1 = yes).
- **vaxperiod**: Indicator variable corresponding to time point, to account for potential decline in antibody response following the post-second dose peak (0 = day 28 or day 28 post second vaccine, 1 = day 90 post second vaccine).
- **ID**: Participant ID.

Table 2.5: Mixed-effects model parameter estimates for $\log_{10}$-transformed antibody responses to dynamic antigens only, modelled from day 28 onwards and stratified by day 28 serostatus. Fixed effects: tp = time in months since second vaccination; vaxperiod = indicator for the post-peak waning period (1 = day 90 post second dose); d28serostatus = MpoxPlex serostatus at day 28 (1 = seropositive). A positive d28serostatus coefficient indicates higher antibody levels in individuals who seroconverted by day 28, independent of the linear time trend. Random effects represent between-participant variability in baseline antibody level (SD Intercept) and residual within-participant variability (SD Residual). Uncertainty intervals and p-values computed using a Wald t-distribution approximation.

Parameter	Coefficient	SE	95% CI	t(81)	p
A35 - Fixed Effects					
(Intercept)	1.58	0.11	[1.37, 1.79]	15.00	$< .001$

Continued on next page

Parameter	Coefficient	SE	95% CI	t(81)	p
tp	0.54	0.07	[0.40, 0.68]	7.66	< .001
vaxperiod	-1.27	0.20	[-1.67, -0.87]	-6.26	< .001
d28serostatus	0.51	0.14	[0.23, 0.80]	3.58	< .001
A35 - Random Effects					
SD (Intercept: ID)	0.36	0.06	[0.26, 0.49]		
SD (Residual)	0.27	0.03	[0.22, 0.33]		
B2 - Fixed Effects					
(Intercept)	1.47	0.05	[1.37, 1.57]	29.69	< .001
tp	0.21	0.03	[0.15, 0.28]	6.22	< .001
vaxperiod	-0.58	0.10	[-0.77, -0.38]	-5.87	< .001
d28serostatus	0.15	0.07	[0.01, 0.28]	2.19	0.031
B2 - Random Effects					
SD (Intercept: ID)	0.16	0.03	[0.12, 0.23]		
SD (Residual)	0.13	0.01	[0.11, 0.16]		
E8 - Fixed Effects					
(Intercept)	2.39	0.09	[2.21, 2.58]	25.67	< .001
tp	0.66	0.07	[0.52, 0.80]	9.31	< .001
vaxperiod	-1.77	0.20	[-2.17, -1.37]	-8.82	< .001
d28serostatus	0.26	0.12	[0.02, 0.51]	2.13	0.036
E8 - Random Effects					
SD (Intercept: ID)	0.30	0.05	[0.21, 0.42]		

Continued on next page

Parameter	Coefficient	SE	95% CI	t(81)	p
SD (Residual)	0.27	0.03	[0.22, 0.32]		
H3 - Fixed Effects					
(Intercept)	1.92	0.09	[1.75, 2.09]	22.25	< .001
tp	0.54	0.07	[0.40, 0.68]	7.46	< .001
vaxperiod	-1.34	0.21	[-1.76, -0.93]	-6.50	< .001
d28serostatus	0.71	0.11	[0.48, 0.93]	6.34	< .001
H3 - Random Effects					
SD (Intercept: ID)	0.26	0.05	[0.18, 0.37]		
SD (Residual)	0.28	0.03	[0.23, 0.33]		
M1 - Fixed Effects					
(Intercept)	1.95	0.12	[1.71, 2.20]	16.00	< .001
tp	0.59	0.06	[0.47, 0.72]	9.66	< .001
vaxperiod	-1.68	0.18	[-2.03, -1.33]	-9.56	< .001
d28serostatus	0.25	0.17	[-0.09, 0.59]	1.45	0.151
M1 - Random Effects					
SD (Intercept: ID)	0.45	0.07	[0.33, 0.60]		
SD (Residual)	0.23	0.02	[0.19, 0.28]		
V.A33 - Fixed Effects					
(Intercept)	1.87	0.10	[1.68, 2.06]	19.28	< .001
tp	0.60	0.07	[0.47, 0.74]	8.69	< .001

Continued on next page

Parameter	Coefficient	SE	95% CI	t(81)	p
vaxperiod	-1.38	0.20	[-1.78, -0.99]	-6.97	< .001
d28serostatus	0.37	0.13	[0.11, 0.63]	2.85	0.005
V.A33 - Random Effects					
SD (Intercept: ID)	0.32	0.05	[0.23, 0.44]		
SD (Residual)	0.27	0.03	[0.22, 0.32]		
V.B5 - Fixed Effects					
(Intercept)	2.32	0.08	[2.17, 2.47]	30.49	< .001
tp	0.60	0.06	[0.48, 0.72]	10.14	< .001
vaxperiod	-1.74	0.17	[-2.08, -1.41]	-10.35	< .001
d28serostatus	0.23	0.10	[0.03, 0.43]	2.28	0.025
V.B5 - Random Effects					
SD (Intercept: ID)	0.24	0.04	[0.17, 0.34]		
SD (Residual)	0.23	0.02	[0.19, 0.27]		

2.4 Discussion

In this prospective cohort study, we detailed the serological and cellular responses to MVA-BN in a predominantly vaccine-naïve cohort receiving fractionated intradermal vaccination. While the primary clinical impetus was to validate dose-sparing strategies during the 2022 mpox outbreak—confirming that fractionated dosing elicits robust seropositivity—our parallel objective was to evaluate MVA-BN as a candidate agent for HIC studies.

2.4.1 MVA-BN as a viral immune challenge model for immunomodulatory drug development

This study supports the utility of MVA-BN as an immune challenge agent for PoM studies of novel IMP. Our data, aligned with numerous prior *in vitro*, animal, and human studies, facilitate a detailed understanding of the interaction between MVA infection and host immunity—a prerequisite for any robust challenge model.

Mechanistically, MVA administration triggers pathways associated with innate viral sensing. This includes the detection of viral genomic DNA in the host cytosol by cyclic GMP-AMP synthase (cGAS), followed by activation of the Stimulator of Interferon Genes (STING). STING activation subsequently induces the expression of Type I interferons, pro-inflammatory cytokines, and Interferon Stimulated Genes (ISGs) such as *IFIH1*, *DDX58*, and *CXCL10*.

Intriguingly, recent work by Yang et al. demonstrated that heat inactivation of MVA (incubation in a water bath at 55°C for 1 hour) resulted in a significantly more potent induction of Type I interferon in mice compared with live virus. This effect is likely due to the inactivation of heat-labile viral immunosuppressive proteins retained in the attenuated live virus, which otherwise dampen the host response. Consequently, heat inactivation represents a simple adjustment to the model that could enhance its utility as a probe of innate and adaptive viral immunity.[30, 31]

2.4.1.1 Differentiation from existing challenge models

We demonstrate that MVA-BN is a potent inducer of Th1-biased immunity and substantial CD8+ T cell responses. This distinguishes MVA-BN from existing protein-adjuvant challenge models, such as Keyhole Limpet Hemocyanin (KLH) or tetanus toxoid, which

predominantly drive CD4+ and humoral responses. Consequently, MVA-BN represents a valuable addition to the experimental medicine toolkit, specifically for testing immunomodulators targeting cytotoxic T-cell pathways or viral sensing mechanisms (e.g., TLR9, cGAS-STING).

Subject to further characterization, an MVA (or heat-inactivated MVA) challenge would have broad applicability to numerous drug classes, particularly in the interrogation of novel therapies for autoinflammatory, autoimmune and degenerative diseases driven by cGAS-STING activation, particularly interferonopathies and Systemic Lupus Erythematosus (SLE) (Table 2.6).

2.4.2 Transcriptomic correlates of immunogenicity and the limits of blood signatures

A major challenge in human in vivo immunology is inter-individual variability. In the publication which accompanies this chapter, we mechanistically interrogated this variability via whole blood transcriptomic analysis (analysis performed primarily by Dr. Nicholas Provine). While this analysis identified distinct baseline gene signatures that correlated with the magnitude of the subsequent humoral response, no equivalent predictive signature was identified for the T cell response.

Detailed presentation of these transcriptomic data is excluded from this chapter as the bioinformatic analysis was performed by collaborators. However, the divergence in these findings—where antibody responses track with peripheral blood transcriptomics but T cell responses do not—is mechanistically revealing. It suggests that the crucial determinants of the cellular response to intradermal vaccination may be localised to the tissue microenvironment (injection site and draining lymph node) rather than the systemic circulation.

This latent immune coordination highlights a fundamental limitation of relying solely on peripheral blood readouts in viral challenge models.

2.4.3 Limitations of the model

The inability to predict cellular responses from blood reinforces the need for tissue-level analysis. The inability to perform skin biopsies in this observational cohort limits our understanding of the local vaccine site micro-environment checkpoint, and should be examined in future experimental medicine studies. Furthermore, reliance on a licensed vaccine product deployed in a pandemic response introduces potential supply chain issues: during the course of the study, MVA-BN supply was severely limited on a global level. Consequently, access was restricted to conduct of the national vaccination programme. While supply constraints may be expected to ease, this highlights the need for a range of experiential challenge agents with uncorrelated supply risks. Finally, while pre-existing immunity in cohorts typically recruited for HIC studies, the historical smallpox vaccination, deployment of this vaccine in the mpox context, and incident mpox disease in common demographics may reduce the pool of potential volunteers who are antigen naive, potentially leading to further heterogeneity in immune responses.

2.5 Conclusion and Future Directions

In summary, MVA-BN represents a safe and scientifically robust viral challenge agent, particularly valuable for interrogating CD8+ T cell biology. Its availability as a GMP-grade commercial product significantly reduces practical and regulatory barriers to adoption compared to novel challenge agents requiring de novo manufacturing, while the intradermal route of administration mirrors established challenge protocols. However, utilising MVA-BN

as an agent for HIC requires further optimisation to establish an optimal (re-)challenge dose, immunological response characteristics, optimal observation timepoints, and sensitivity to pharmacological modulation. Our study highlights that peripheral blood markers alone are insufficient to capture the full cellular immune landscape; future development must therefore focus on granular tissue sampling—specifically of skin biopsy sites and draining lymph nodes—to accurately map the pharmacodynamic effects of candidate immunomodulators.

In the context of the thesis, to advance work towards better experimental and analytical paradigms for immune challenge, we chose to focus subsequently on another agent with a pedigree of existing use in the HIC contex: KLH. As a neoantigen independent of clinical supply pressures, KLH offers a tractable and experimentally flexible system to optimise the tissue-centric protocols required for enhanced HIC models.

Table 2.6: Immunomodulatory drug classes potentially amenable to MVA-HIC PoM assessment

Drug Class	Example drug(s)	Target Pathway / Mechanism	Target Disease(s)
cGAS Inhibitors	VENT-03 (Ventus Therapeutics) IMSB301 (ImmuneSensor Therapeutics)[32, 33]	**Cytosolic DNA Sensing:** Prevents cGAS from binding dsDNA and synthesizing cGAMP.	Monogenic autoinflammatory syndromes (e.g. Aicardi-Goutières syndrome), SLE, Rheumatoid arthritis. Also potential relevance for diverse neurological, metabolic, cardiovascular and inflammaging-related disorders [34]
STING Antagonists	H-151 [34] Palbociclib [35]	**STING Signaling:** Blocks STING palmitoylation or dimerization, preventing downstream IFN-I induction.	Similar to target diseases for cGAS ingibitors with rare exceptions including STING-Associated Vasculopathy with onset in Infancy and COPA.
pDC Modulators	Litifilimab (anti-BDCA2)[36] Daxdilimab (anti-ILT7)[37]	**pDC Depletion/Inhibition:** mAbs that sequester or deplete Plasmacytoid DCs to halt IFN-α production.	SLE, Cutaneous Lupus
TLR7/9 Antagonists	Enpatoran Afimetoran	**Endosomal Sensing:** Blocks detection of ssRNA/dsDNA in endosomes; synergistic with STING blockade.	SLE, Sjögren's Syndrome
TYK2 Inhibitors	Deucravacitinib NDI-034858	**Cytokine Signaling:** Selectively blocks the JAK-STAT arm responsible for IL-23, IL-12, and Type I IFN signaling.	Psoriasis, SLE

Chapter 3

In vivo human keyhole limpet hemocyanin challenge in early phase drug development: A systematic review

Chapter summary

Previous reviews of keyhole limpet haemocyanin (KLH) immune challenge studies have identified important methodological deficiencies; however, it has remained unclear whether these limitations persist in contemporary studies, particularly in the context of early-phase drug development. Moreover, no prior reviews have critically evaluated the application of KLH challenge as a human immune challenge model for proof-of-mechanism assessment of immunomodulatory investigational medicinal products (IMPs).

This chapter presents a systematic review of KLH immune challenge studies, with a specific focus on their use in early-phase clinical development.

This work addresses Objective 2 of the thesis:

To systematically identify, collate, and synthesise HIC studies employing KLH, in order to define design limitations and inform a research agenda for optimisation of KLH HIC paradigms in drug development.

This work has been published in *Clinical and Translational Science* (DOI: 10.1111/cts.13457).[38]

The review demonstrates that, while KLH challenge is widely used to interrogate immune engagement and mechanism in humans, its application is highly heterogeneous, with substantial variability in challenge protocols, immunological endpoints, and reporting standards. Notably, relatively few studies incorporate tissue-level or site-of-priming immune readouts, limiting the interpretability of KLH challenge as a robust proof-of-mechanism tool.

3.1 Introduction

Whilst many antigens may be employed in the conduct of HIC studies, the xenogenic neoantigen Keyhole Limpet Haemocyanin (KLH, derived from the Grand Keyhole Limpet *Megathura crenulata*) is considered a 'model antigen' for this purpose.[16] Because humans are commonly naïve to KLH epitopes prior to immunisation (albeit with cross-reactive responses to structurally similar endogenous or exogenous carbohydrate epitopes in some cases).[39, 40] KLH has advantages over alternative natural or therapeutically employed antigens commonly used for HIC (e.g. varicella zoster or BCG) which do not allow control of the degree, duration, and time-elapsed since primary exposure. KLH has been used for over 50 years for TDAR/DTH assessment in both HIC and preclinical contexts.[41–44]. Long-standing use has affirmed its utility as a safe and reliable model inducer of a TDAR/DTH response. However, despite widespread use in interventional and observational studies, HIC employing KLH ('KLH challenge') has historically lacked fundamental characterisation and standardisation.[16, 40] For example, early KLH preparation methods yielded a product

which contained substantial impurities, which could result in unpredictable effects on the immune response, and deleterious impacts on internal and external study validity. KLH is now available as a highly-purified, GMP-grade product, as either a 'native' or high molecular weight (HMW) form, or a 'sub-unit' form consisting of disaggregated sub-units (designated KLH1 and KLH2) of approximately 400kDa each.[40] Although early studies have explored the dose-response of KLH neoantigen and DTH responses, to our knowledge no studies in the modern era have systematically explored antigen dose-response using GMP-grade (sub-unit or HMW) KLH.[45–48]

Given the potential value of KLH challenge in immunomodulatory drug development, and previous concerns regarding its optimisation, we sought to describe the current status of KLH challenge in this context. The aim of this review was therefore to systematically identify, collate and describe HIC studies employing KLH.

3.1.1 Objectives

The objectives of this chapter were:

Chapter Objectives

1. To identify studies applying KLH challenge, focusing on application to early phase clinical trials, and related studies which may inform the optimisation of the model

2. To describe applications of human KLH challenge in terms of KLH formulation, use of adjuvants, dose regimen, outcomes assessed, and concordance of outcomes with preclinical immune challenge studies (employing KLH and other T-cell dependent antigens)

3. To identify opportunities for further refinement of the model, in order to maximise the utility of this approach in early phase clinical trials of immunomodulatory IMP.

3.2 Methods

We used the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework to structure this review, and managed the review process using the Rayyan online systematic review platform.[49, 50]

3.2.1 Search strategy

We identified studies conducted between 1 January 1994 and 1 April 2022, where KLH was administered to humans via percutaneous injection to elicit an immune response for non-therapeutic purposes.

We searched the following databases and registries: i) Medline (Ovid), ii) EMBASE (Ovid), iii) clinicaltrials.gov, iv) Cochrane CENTRAL trials register. The full electronic search

strategy is available in table 3.1. Reference lists of included studies and previous relevant reviews were also searched.[16, 40] Initial screening of titles and abstracts was performed by a single reviewer, followed by independent review of potentially relevant manuscripts by two reviewers. Any conflicts were resolved following review of the manuscript and discussion between the reviewers. Following Swaminathan et al., we restricted our search to studies published after 1994, due to the significant uncertainty associated with generalisation of earlier studies using poorly purified, non-GMP grade formulations of KLH.

3.2.2 Inclusion and exclusion criteria

We included published studies where KLH was administered to healthy volunteers or patient populations via percutaneous routes (i.e. intramuscular, subcutaneous, or intradermal injection). The following studies were excluded: i) studies where KLH was administered as a vaccine adjuvant or control substance in cancer vaccine studies, ii) in vitro or ex vivo analyses of human tissue or cells where KLH was not administered to participants prior to sampling, iii) studies where KLH was administered as part of a therapeutic regimen (e.g. bladder cancer), iv) studies where KLH was given by non-percutaneous routes (e.g. enteral, inhalational), and v) studies reported as conference abstracts only.

3.2.3 Reporting quality

We assessed reporting quality using the Template for Intervention Description and Replication (TIDieR) checklist, adapted to capture key aspects relevant to the reproducible conduct of KLH challenge (table 3.2).[51] Given the intended focus of the review, the reporting quality assessment was restricted to the studies applying KLH challenge to early phase clinical trials. Appraisals of the two independent reviews were compared and any disagreements

resolved through discussion and consensus, or by third party adjudication if necessary.

3.2.4 Data extraction

We categorised identified studies into: (i) early phase (1 or 2) clinical trials, (ii) late phase clinical trials, post-marketing surveillance, or observational studies of immunomodulatory drugs, (iii) other contexts e.g., exploration of fundamental immunology in healthy participants, effect of a disease state, or effect of a non-pharmacological exposure or intervention. Variables extracted from each study included date of publication, study aim, exposures (both pharmacological and non-pharmacological), KLH regimen (product, source, formulation, use of adjuvant, dose, timing of rechallenge dose(s)), and outcomes assessed (KLH-specific and non-specific pharmacodynamic endpoints). For IMPs, the rationale for a specific KLH challenge regimen (e.g. KLH dose, timing relative to the IMP) and evidence of modulation of KLH-specific pharmacodynamic endpoints or PoM was described, where reported. Finally, to further understand the translational relevance of the KLH challenge approach to early phase studies of IMP, where a corresponding T-cell dependent antigen challenge-drug combination was reported by the study authors, we evaluated the concordance of the findings between the preclinical and clinical studies.

Table 3.1: KLH systematic review search strategy

Database	Search Details
ClinicalTrials.gov	
	• Search: KLH OR "keyhole limpet hemocyanin" = 150 trials
	• Records selected for further review: 25

Database	Search Details
Cochrane Central	
	• Search date: 30 April 2022 • Search: "KLH" or "Keyhole Limpet" = 258 trials • Records selected for further review: 44
OVID Medline	
	• Search date: 30 April 2022 • Database: Medline (Ovid MEDLINE® Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE® Daily, and Ovid MEDLINE®) 1946 to present • Search steps: 1. KLH.mp. = 2970 2. keyhole limpet hemocyanin.mp. = 3871 3. keyhole limpet haemocyanin.mp. = 504 4. 1 or 2 or 3 = 4918 5. exp animals/ not humans.sh. = 5001776 6. mice.mp. or Mice/ = 1847098 7. bladder cancer.mp. or Urinary Bladder Neoplasms/ = 69969 8. 5 or 6 or 7 = 5789490 9. 4 not 8 = 1145 10. limit 9 to yr="1994 -Current" = 847 11. from 10 keep 16, 23, 40-41, 44, 50, 56, 58, 63, 84, 103, 109, 138, 161-162, 170, 189, 223, 226, 274, 292, 298, 303, 326, 353, 432, 478, 481, 486, 491, 502, 520, 558, 566, 660, 664, 675, 679, 694, 751, 755, 776, 780, 816 • Records selected for further review: 44

Database	Search Details
OVID EMBASE	• Search date: 30 April 2022 • Database: Embase 1974 to present • Search steps: 1. keyhole limpet hemocyanin/ = 3365 2. KLH.mp. = 3743 3. 1 or 2 = 5402 4. exp human/ or human tissue.mp. = 23574384 5. exp animal/ or exp invertebrate/ or nonhuman/ or animal experiment/ or animal tissue/ or animal model/ or exp plant/ or exp fungus/ = 30903717 6. 5 not 4 = 7343949 7. 3 not 6 = 2229 8. cancer vaccine/ or tumor vaccine/ or lymphoma vaccine/ or dendritic cell vaccine/ = 23618 9. bladder cancer/ = 54414 10. 8 or 9 = 77630 11. 7 not 10 = 1644 12. limit 11 to yr="1994 -Current" = 1344 13. in vitro study/ = 1411012 14. 12 not 13 = 1220 15. malignant neoplasm/ = 77761 16. 14 not 15 = 1216 17. lymphoma/ = 110156 18. 16 not 17 = 1197 19. from 19 keep 2, 21-22, 34, 41, 44, 60, 70-71, 87-88, 97, 99, 119, 122, 127, 131, 137, 142, 155-156, 190, 197, 210-211, 219, 227, 235, 237, 241, 249, 291-292, 305, 331, 352, 371-372, 399, 404, 424, 514, 579, 589, 595, 647, 689, 792, 809, 861, 876, 878, 912, 946, 997, 1083-1084, 1139 • Records selected for further review: 58

Table 3.2: Modified TIDieR reporting checklist for early phase clinical trials of IMP employing KLH challenge

Checklist component	Guidance notes
2a. Describes rationale for using KLH challenge	Minimum requirement is some specificity. E.g. 'to examine PD effect' would be insufficient, but 'to examine effects on neoantigen response' would be sufficient.
2a: detail	What was the rationale for using KLH challenge.
2b. Describes rationale for a specific challenge regimen	Score yes if any discussion of rationale (e.g. why specific endpoint or timings chosen).
2b: detail	What was the rationale for the specific challenge regimen.
3a. Describes KLH form	
3b. Describes KLH product name	
3c. Describes KLH supplier/company	
3d. Describes adjuvant generic name	
Describes adjuvant product name	
3f. Describes adjuvant supplier	
4a. Describes KLH preparation procedure	
4b. Describes KLH route of administration	
4c. Describes KLH site of administration	
4d. Describes rechallenge route of administration	
4e. Describes rechallenge site of administration	
4f. Describes KLH administration device	E.g. '27g needle'.
5a. Describes who administered the challenge agent	E.g. administered by trained study staff.
7. Describes where the study was conducted	
8a. Describes primary KLH dose	
8b. Describes primary KLH timepoint relative to other study interventions	
8c. Describes rechallenge KLH doses	
8d. Describes rechallenge KLH timepoints	
8e. Describes adjuvant dose	

Checklist component	Guidance notes
9. Describes use of KLH challenge in title	
10a. Describes use of KLH challenge in abstract	
10b. Refers to assessment of immune response of an administered antigen in abstract	
Other: Blinding: were assessors of KLH DTH clinical response blinded to dose level or placebo	NA if no placebo, single dose level, or if no DTH performed.

Table 3.3: Early phase (1 and 2) clinical trials (n=14) of investigational medicinal products employing KLH challenge

Study Population (Number)	KLH Regimen: Formulation, Dose, Route, Site, Timing	Antibody Assessment (Method, Timepoint)	Other KLH-Specific PD Endpoints	Non-KLH Specific PD Endpoints	Evidence for Modulation of KLH-Specific Endpoints / Proof-of-Mechanism
Abatacept: CTLA4-Ig fusion protein, antagonist of the CD28-CD80/86 interaction (Abrams, 1999, Phase 1)[52]					
Patients with psoriasis (n=43)	HMW-KLH (Intracel), 1000 µg ID, administered 14 days before and 29 days after initiation of abatacept or control	Anti-KLH Ig total (ELISA), measured 14 and 28 days after second KLH immunisation	nil	Bacteriophage ϕX174 neoantigen response, Psoriasis disease activity, Serial histological response of psoriatic lesions, Blood lymphocyte subpopulations	Suppression of antibody titres observed in 1 or more patients accrued to all dose levels. Normal peak antibody titres observed in some patients in the highest dose cohort. Qualitatively similar observations in preclinical models.
Rituximab: chimeric anti-CD20 mAb (Van der Kolk, 2002, Phase 1/2)[53]					
Patients with relapsed low-grade lymphoma (n=11 total, n=7 received KLH challenge)	HMW-KLH (Calbiochem) 1000 µg SC, either 14 days before or 14 days after rituximab treatment	Anti-KLH IgG (ELISA) assessed 14 days post-KLH	nil	Antibody responses to Hepatitis A vaccine, tetanus toxoid, and poliomyelitis vaccine, Total immunoglobulin concentrations, Peripheral B cell count	None of the patients developed anti-KLH IgG responses before or after rituximab treatment.
Rituximab (Bingham, 2010, Phase 2)[54]					
Patients with rheumatoid arthritis receiving methotrexate (n=103)	HMW-KLH (Intracel) 1000 µg SC, 36 weeks post rituximab course	Anti-KLH IgG at baseline and 28 days post KLH	nil	Antibody responses to pneumococcal polysaccharide antigen and tetanus toxoid, Total immunoglobulin levels, Lymphocyte subsets	Proportion of patients mounting a quantifiable anti-KLH IgG response was lower in the rituximab-treated group (47%) than in the MTX alone-treated group (93%).
Recombinant human HGH (Somatotrophin) (Smith, 2010, Phase 2)[55]					
Patients with HIV infection (n=60)	Sub-unit KLH (Immucothel), 100 µg ID, 16 and 20 weeks post initiation of rhGH or standard care	nil	Lymphoproliferative responses to in vitro KLH stimulation of PBMCs	HIV-1 viral load, T-lymphocyte subsets, PBMC lymphoproliferative assays, T cell receptor circles, thymus size on CT scan	Two participants in the intervention arm (n=30), and no participants in the control arm (n=30) developed new lymphoproliferative responses to KLH at week 24.
Recombinant CD40L (Jain, 2011, Phase 2)					
Paediatric patients with X-linked hyper IgM syndrome (n=3)	KLH (formulation, source and route not stated) 2500 µg, 8, 12, and 16 weeks after initiation of regular rCD40L. Three subsequent ID KLH challenges (dose not stated)	Anti-KLH IgG (ELISA)	DTH response (induration)	DTH response to recall antigens, Serum total immunoglobulins, Antibody response to ϕX-174, Cytokine release, PBMC stimulation, B cell mutation analysis, CT imaging of axillary lymph nodes, Histology on lymph node biopsy	All 3 participants developed a positive DTH response during rCD40L treatment, followed by negative responses after a 12-week drug-free interval. No patients developed detectable anti-KLH IgG.
MEDI6469/9B12: OX40 agonist (Curti, 2013, Phase 1)[56]					
Patients with advanced solid malignancies (n=30)	Sub-unit KLH (Immucothel) SC (dose not stated). Arm A: KLH day 1, arm B: day 29, relative to anti-OX40 initiation	Anti-KLH IgG (ELISA) assessed day 15 (arm A) or day 43 (arm B)	nil	Anti-tetanus antibody, Blood lymphocyte subpopulations, Tumour-specific T cell assays, T cell proliferation, Radiological tumour response	A greater increase in KLH antibody titre was observed in patients immunised on the same day as anti-OX40 compared to those immunised 28 days later.
FR104: pegylated anti-CD28 monovalent Fab' antibody (Poirier, 2016, Phase 1)[57]					
Healthy volunteers (n=64 total, n=33 in KLH challenge cohorts)	KLH (formulation, source, route, dose not stated—likely Immucothel), presumed given on the same day as FR104 (SAD cohorts)	Anti-KLH IgG (ELISA) assessed at screening, day 15, 29, 57, 85, and/or 113	nil	Serum cytokine concentration, CD28 receptor occupancy, T lymphocyte subsets in blood, Ex vivo SEB and LPS stimulation, EBV reactivation	Dose-dependent inhibition of time to detection of anti-KLH IgG.
Lulizumab pegol: BMS-931699 (Shi, 2016, Phase 1)[58]					
Healthy volunteers (n=16)	Sub-unit KLH (Immucothel) 1000 µg SC given on the same day as lulizumab	Anti-KLH IgM and IgG (ELISA), assessed day 1, 8, 15, 29	nil	Serum cytokine concentration, Lymphocyte subsets, Receptor occupancy on CD4+ and CD8+CD11a$_{low}$ T lymphocytes	Higher doses of lulizumab resulted in inhibition of anti-KLH IgG titres for $\geq$2 weeks post-dose. Suppression associated with > 80% CD28 receptor occupancy.
Prezalumab (AMG557): human ICOSL mAb (Sullivan, 2016, Phase 1)[59]					
Patients with systemic lupus erythematosus (n=112)	Sub-unit KLH (Vacmune) 1000 µg ID, various timings post-prezalumab	Anti-KLH IgM and IgG (flow cytometric bead array)	nil	ICOSL target occupancy, Total and free ICOSL levels on B cells, Tetanus antitoxin IgG, SLE biomarkers	Decreased anti-KLH IgG in MAD cohorts. Evidence of dose response correlating with IC99 target receptor occupancy.

Table 3.3: (Continued)

Study Population (Number)	KLH Regimen: Formulation, Dose, Route, Site, Timing	Antibody Assessment (Method, Timepoint)	Other KLH-Specific PD Endpoints	Non-KLH Specific PD Endpoints	Evidence for Modulation of KLH-Specific Endpoints / Proof-of-Mechanism
Baminercept: lymphotoxin-β Receptor IgG Fusion Protein (St Clair, 2018, Phase 2)[60]					
Patients with Sjogren's syndrome (n=52)	Sub-unit KLH (Immucothel) 1000 µg SC with Montanide ISA-51 VG, 8 weeks post-initiation of weekly baminercept	Anti-KLH IgM and IgG (ELISA), assessed day 28 post-KLH	nil	Pneumococcal polysaccharide vaccine antibody response, Sjogren's syndrome clinical endpoints	Baminercept did not suppress the anti-KLH IgG response. Increase in anti-KLH IgM reported.
Karnell (2019), Phase 1: Dazodalibep (VIB4920/ MEDI4920/ HZN4920)[61]					
Healthy volunteers (n=56)	KLH (formulation and source not stated) plus aluminium hydroxide 1000 µg SC 14 days before and 15 days post dazodalibep	Anti-KLH IgG (ELISA) day 22, 29, 43, 57, 85, and 113 post dazodalibep	nil	sCD40L as a measure of target engagement	Dose-dependent inhibition of anti-KLH IgG. No inhibition at low doses, near-complete suppression at the highest dose level. Qualitatively similar response in mice immunised with SRBC.
Espié (2019), Phase 1: Iscalimab (CFZ533)[62]					
Healthy volunteers (n=36)	KLH (formulation and source not stated) 115 µg IM day 3 post-iscalimab, second KLH dose at varying timepoints corresponding to expected CD40 receptor occupancy loss (day 29, 43, 57, 71, 82, 85)	Anti-KLH IgM and IgG (assessed at day 8, 15, 22 post-first KLH dose, and 14, 28, 42 post-second dose)	nil	Free/total CD40 receptor levels (whole blood B cells), EBV/CMV reactivation, serum cytokine concentrations	Higher iscalimab doses transiently suppressed anti-KLH responses for approximately 3–4 weeks. All subjects mounted responses to the rechallenge KLH dose. Observations aligned with non-human primate studies.
Yang (2021), Phase 1: Acazicolcept (ALPN-101)[63]					
Healthy volunteers (n=92 total, n=36 in KLH challenge cohorts)	KLH (Stellar Biotechnologies, formulation not stated) 1000 µgSC day 0 or 1 post-acazicolcept administration	Anti-KLH IgM and IgG (ELISA), assessed days 0, 7, 14, 21, 28 in SAD cohort, and additional timepoints in MAD cohort	nil	Serum cytokine concentration, CD4/CD8 T-cell target saturation (CD28 and ICOS), ex vivo SEB stimulation pre/post-dose, cytokine quantification	Single and repeated doses of acazicolcept significantly reduced anti-KLH antibody titres compared to placebo. Trends toward dose-dependent inhibition observed in SAD cohorts. Concordance noted with mouse studies on KLH and SRBC.
Saghari (2022), Phase 1: KY1005 (OX40L Antagonist)[64]					
Healthy volunteers (n=64)	Sub-unit KLH (Immucothel) 100 µg with aluminium hydroxide 900 µg IM, day 7 post-third dose of KY1005 (administered 0, 4, and 8 weeks). ID rechallenge KLH 1 µg day 21 post-primary dose	Anti-KLH IgM and IgG (ELISA) day 21 post-primary KLH immunisation	DTH response (induration, erythema)	OX40 and OX40L expression on PBMCs, anti-tetanus toxoid antibody titres	No significant difference in antibody titres across dose levels. Post hoc analysis suggested modest exposure-response relationships for IgM and IgG. Observations partially concordant with non-human primates.

Table 3.4: Late phase and post-marketing trials (n=6) investigating drug effects on response to KLH challenge

Drug: Mechanism	Study Population (Number)	KLH Regimen	Antibody Assessment (Method, Timepoint)	Other KLH-Specific PD Endpoints	Evidence for Modulation of KLH-Specific Endpoints
Rentenaar (2002)[65], observational: Three cohorts (P/CsA; P/CsA/anti-CD3 mAb; P/CsA/MMF)					
Prednisolone/Cyclosporine A or combined regimens	Patients with renal transplant (n=34)	HMW-KLH (Calbiochem) 1000 µg SC, rechallenge 100 µg ID day 14	Anti-KLH IgG (ELISA), day 14 post-KLH	DTH response: induration, cell markers (CD3, ICAM-1, VCAM-1)	Significant reduction in anti-KLH IgG in transplant patients compared to healthy volunteers. All healthy controls developed DTH, while only two transplant groups showed responses.
Saville (2008)[66], phase 4/post-marketing surveillance: Rituximab					
Rituximab	Healthy volunteers (n=84), NHL patients (n=110)	KLH (formulation/source not stated), 36 weeks post-rituximab (weekly x4 doses)	Anti-KLH IgG (method not stated), day 28 post-KLH	Pre-existing bacterial/viral titres, antibody response to tetanus toxoid	4/108 NHL patients had doubling in anti-KLH titre vs. 69% of healthy volunteers.
Struijk (2010)[67], observational: Three cohorts (P/CsA; P/mycophenolic acid; P/everolimus)					
P/CsA, P/mycophenolic acid, or P/everolimus	Healthy volunteers (n=13), renal transplant patients (n=36)	Sub-unit KLH (Immucothel) 1000 µg SC	Anti-KLH IgG (ELISA) day 14 post-KLH	Antibody responses to pneumococcal polysaccharide and tetanus toxoid; ELISPOT for IL-2, IFNγ, IL-4	Anti-KLH IgG response lower in patient cohorts vs. healthy controls. Evidence of increased IgG in everolimus group.
Boulton (2012)[68]: Fingolimod (sphinosine-1-phosphate receptor modulator)					
Fingolimod	Healthy volunteers (n=72)	Sub-unit KLH (Immucothel) 100 µg + aluminium hydroxide IM day 7, 14, 21; rechallenge KLH 10 µg day 28	Anti-KLH IgM/IgG (baseline, day 7, 14, 21, 28, 42, 56)	DTH response to tetanus and candida antigens, lymphocyte subsets	Lower IgM/IgG in fingolimod groups, with dose-response trend at higher doses.
Kaufman (2014)[69], phase 4/post-marketing surveillance: Natalizumab (humanised IgG4κ mAb)					
Natalizumab	MS patients (n=60)	Sub-unit KLH (Immucothel) 1000 µg SC day 0, 14, 28 (2 or 6 months post-natalizumab)	Anti-KLH IgG, day 14, 28, 56	Antibody response to tetanus toxoid	No discernible effect on anti-KLH IgG responses.
Bar-Or (2020)[70], phase 3b: Ocrelizumab (humanised anti-CD20 mAb)					
Ocrelizumab	MS patients (n=102)	KLH (source not stated) 1000 µg IM day 0, 28, 56 (12 weeks post-ocrelizumab dose)	Anti-KLH IgM/IgG, day 28, 56, 84	Antibody responses to pneumococcal polysaccharide/conjugate vaccine, influenza vaccine, tetanus toxoid; PB B cell counts	Significant reduction in IgM/IgG titres in ocrelizumab group compared to control at 4, 8, and 12 weeks.

Table 3.5: Other human studies employing KLH challenge (n=26)

Lead Author	Year Published	Title	Aim
Kondratenko[71]	1997	Lack of specific antibody response in common variable immunodeficiency (CVID) associated with failure in production of antigen-specific memory T cells. MRC Immunodeficiency Group	To evaluate responses to primary KLH immunization in patients with immunodeficiency states.
Kuijpers[72]	1997	Antigen-specific immune responsiveness and lymphocyte recruitment in leukocyte adhesion deficiency type II	To explore the in vivo immune responsiveness and lymphocyte recruitment to the skin, in response to the neo-antigen keyhole limpet hemocyanin (KLH) in a LAD-II patient.
Suchin[73]	1999	Extracorporeal photochemotherapy does not suppress T- or B-cell responses to novel or recall antigens	To determine the effect of photopheresis on humoral and cell-mediated immunity in human subjects.
Valdez[74]	2000	Response to immunization with recall and neoantigens after prolonged administration of an HIV-1 protease inhibitor-containing regimen. ACTG 375 team. AIDS Clinical Trials Group	To assess response to immunization after prolonged anti-retroviral therapy in patients with HIV.
Markert[75]	2000	Effect of highly active antiretroviral therapy and thymic transplantation on immunoreconstitution in HIV infection	To address whether thymus transplantation in concert with HAART would facilitate the regeneration of the immune system in HIV infection.
Lange[76]	2003	Nadir CD4+ T-cell count and numbers of CD28+ CD4+ T-cells predict functional responses to immunizations in chronic HIV-1 infection	To ascertain whether delaying the initiation of highly active antiretroviral therapy compromises functional immune constitution in HIV-1 infection in persons who regain 'normal' CD4 T-cell counts after suppressive antiretroviral therapies.
Lange[77]	2003	Proliferation responses to HIVp24 during antiretroviral therapy do not reflect improved immune phenotype or function	To ascertain whether lymphoproliferation (LP) responses to HIVp24 in chronically infected patients treated with ART predict and improved cytolytic T-cell phenotype or better in vivo immune function as measured by immunisation responses.
Boelens[78]	2003	Glutamine-enriched enteral nutrition increases in vitro interferon-gamma production but does not influence the in vivo specific antibody response to KLH after severe trauma. A prospective, double blind, randomised clinical study	To compare the effect of glutamine-enriched enteral nutrition with control enteral feeding on type I and type II T-lymphocyte responses in patients with severe trauma.
Smith[79]	2004	Influence of age and physical activity on the primary in vivo antibody and T cell-mediated responses in men	To examine the effect of age and physical activity on primary immune response to KLH immunization.
Boelens[80]	2004	Primary immune response to keyhole limpet haemocyanin following trauma in relation to low plasma glutamine	To examine the effect of severe trauma on early primary immune response to KLH immunization in relation to low plasma glutathione.
Smith[81]	2004	The relationship between distress and the development of a primary immune response to a novel antigen	To examine the effect of distress on primary KLH immunization response in young adults.
Smith[82]	2004	Influences of distress and alcohol consumption on the development of a delayed-type hypersensitivity skin test response	To examine the effect of psychological distress on DTH response following primary KLH immunization in young adults.
Miller[83]	2005	Diminished neo-antigen response to keyhole limpet hemocyanin (KLH) vaccines in patients after treatment with chemotherapy or hematopoietic cell transplantation	To compare the responses to KLH immunization in healthy adults with those in immunosuppressed patients (cancer and bone marrow transplant recipients).
Grant[84]	2008	Cardiovascular exercise intervention improves the primary antibody response to keyhole limpet hemocyanin (KLH) in previously sedentary older adults	To examine the effect of aerobic exercise in sedentary older adults on primary immune response to KLH immunization.
Spazierer[85]	2009	T helper 2 biased de novo immune response to Keyhole Limpet Hemocyanin in humans	To establish an immunization protocol to induce de novo Th2 responses using immunization with KLH.
Oyelaran[39]	2010	Evaluation of human antibody responses to keyhole limpet hemocyanin on a carbohydrate microarray	To evaluate antibody responses to KLH in humans using a novel carbohydrate microarray.
Rodriguez[86]	2010	In vitro naive T cell proliferation failure predicts poor post-immunization responses to neoantigen, but not recall antigens, in HIV-infection	To investigate whether naive T cell proliferation could predict in vivo responses to immunization in HIV.

Lead Author	Year Published	Title	Aim
Milgrom[87]	2012	Response to cutaneous immunization with low-molecular-weight subunit keyhole limpet hemocyanin	To determine whether humoral and cellular immune responses would be provoked by cutaneous administration of keyhole limpet hemocyanin and in particular by scarification of the skin.
Ferbas[88]	2013	A novel assay to measure B cell responses to keyhole limpet haemocyanin vaccination in healthy volunteers and subjects with systemic lupus erythematosus	To assess performance characteristics of immunoassays measuring antigen specific response to KLH immunization in healthy controls and patients with systemic lupus erythematosus.
Gallegos[89]	2013	Toward identifying the effects of the specific components of Mindfulness-Based Stress Reduction on biologic and emotional outcomes among older adults	To examine the effects of mindfulness-based stress reduction on immunological outcomes in older adults.
Moynihan[90]	2013	Mindfulness-Based Stress Reduction for Older Adults: Effects on Executive Function, Frontal Alpha Asymmetry and Immune Function	To explore the effect of mindfulness-based stress reduction for older adults on executive function, left frontal asymmetry and antibody response.
Belson[91]	2016	Characterisation of the clinical and activated T cell response to repeat delayed-type hypersensitivity skin challenges in human subjects, with KLH and PPD, as a potential model to test T cell-targeted therapies	To characterise the delayed-type hypersensitivity (DTH) skin reaction to repeated challenges of keyhole limpet hemocyanin (KLH) and tuberculin purified protein derivative (PPD) in healthy volunteers, as a potential model to test T cell-targeted investigational agents.
Giesecke[92]	2018	Simultaneous Presence of Non- and Highly Mutated Keyhole Limpet Hemocyanin (KLH)-Specific Plasmablasts Early after Primary KLH Immunization Suggests Cross-Reactive Memory B Cell Activation	To describe the progression of human primary and secondary humoral immunity following KLH administration.
Swaminathan[93]	2019	Exposure to Solar UVR Suppresses Cell-Mediated Immunization Responses in Humans: The Australian Ultraviolet Radiation and Immunity Study	To assess the influence of natural sun exposure on the sensitization phase of a primary immune response to a model protein antigen, keyhole limpet hemocyanin.
Saghari[94]	2020	A randomised controlled trial with a delayed-type hypersensitivity model using keyhole limpet haemocyanin to evaluate adaptive immune responses in man	To objectively quantify KLH-specific DTH responses, in relation to KLH-specific circulating antibody responses, using skin blood perfusion measurements.
Otterhaug[95]	2021	Photochemical Internalization Enhanced Vaccination Is Safe, and Gives Promising Cellular Immune Responses to an HPV Peptide-Based Vaccine in a Phase I Clinical Study in Healthy Volunteers	To assess the safety and local tolerance of phytochemical internalisation mediated vaccination.

3.3 Results

3.3.1 Human KLH challenge applications: 1994-2022

We identified 46 studies published between 1994-2022 which met the inclusion criteria (figure 3.1). 14 studies (30%) described applications to early phase drug development (table 3.3), and six (13%) described applications to late phase clinical trials, post-marketing surveillance, or observational studies of immunomodulatory drugs (table 3.4). 26 studies (57%) described application of the model for other purposes, including five (11%) describing responses in healthy volunteers (including model optimisation prior to intended application in a clinical trial [91, 94, 96]), 11 (24%) describing effects of non-pharmacological exposures, seven (15%) describing effects of disease states, two (4%) describing development of novel assays of KLH-specific responses and one (2%) describing the assessment of a novel adjuvant table 3.5). An increasing number of studies applied KLH challenge for the assessment of immunomodulatory drugs over the review period (figure 3.2): between 1994-2008, four studies described KLH challenge in the context of pharmacological exposure (including two early phase clinical trials), compared to 16 studies between 2009-2022 (12 early phase trials).

3.3.2 Early phase (I or II) clinical trials

Of the 14 early phase clinical trials identified, six (43%) recruited healthy volunteers, and eight (57%) recruited patients with autoimmune conditions (psoriasis,[52] rheumatoid arthritis,[54] systemic lupus erythematosus,[59] Sjogren's syndrome [60]), inherited immunodeficiencies,[97] HIV,[55] or malignancy.[53]

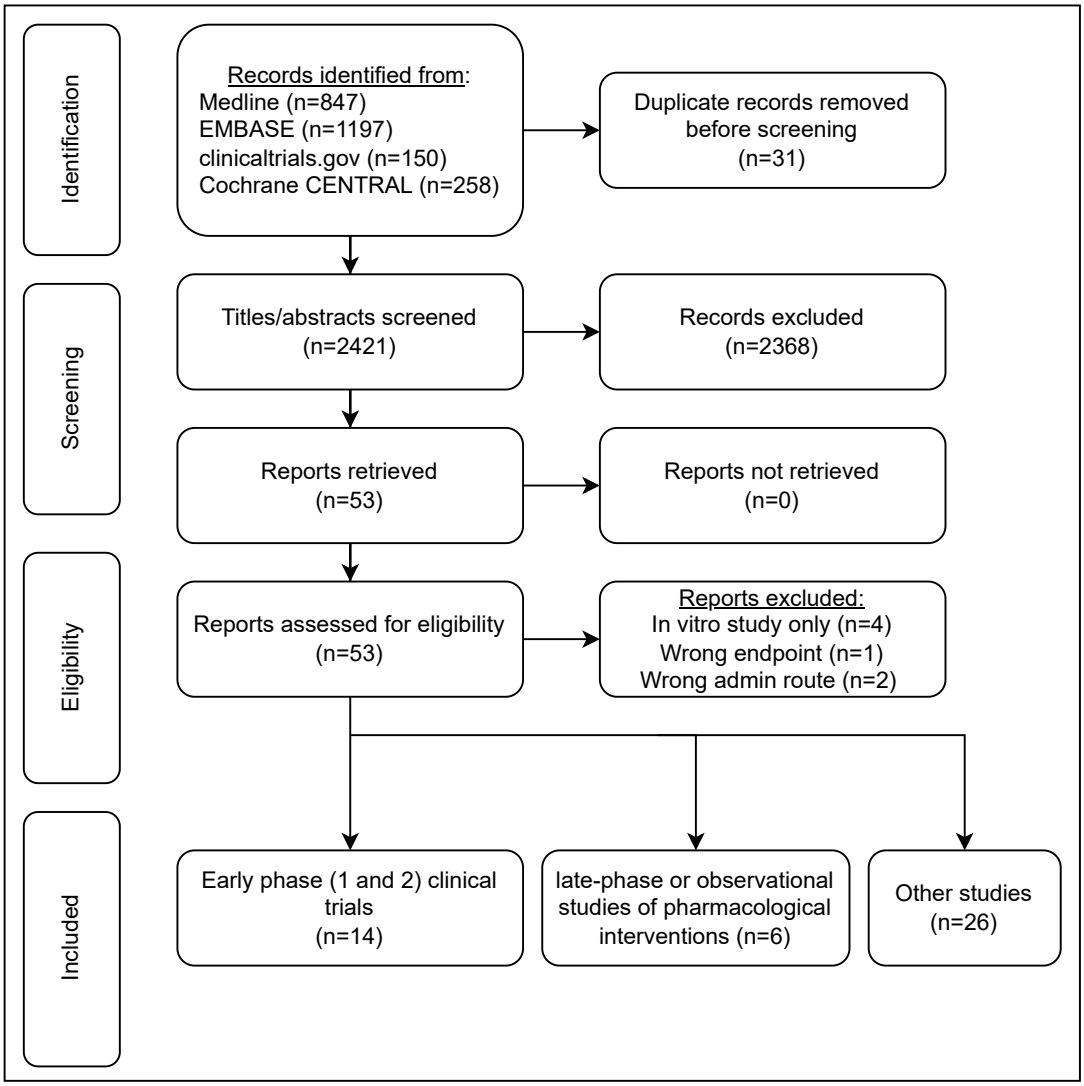

Figure 3.1: Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram describing systematic literature search for experimental HIC studies employing KLH. Image reproduced from [38].

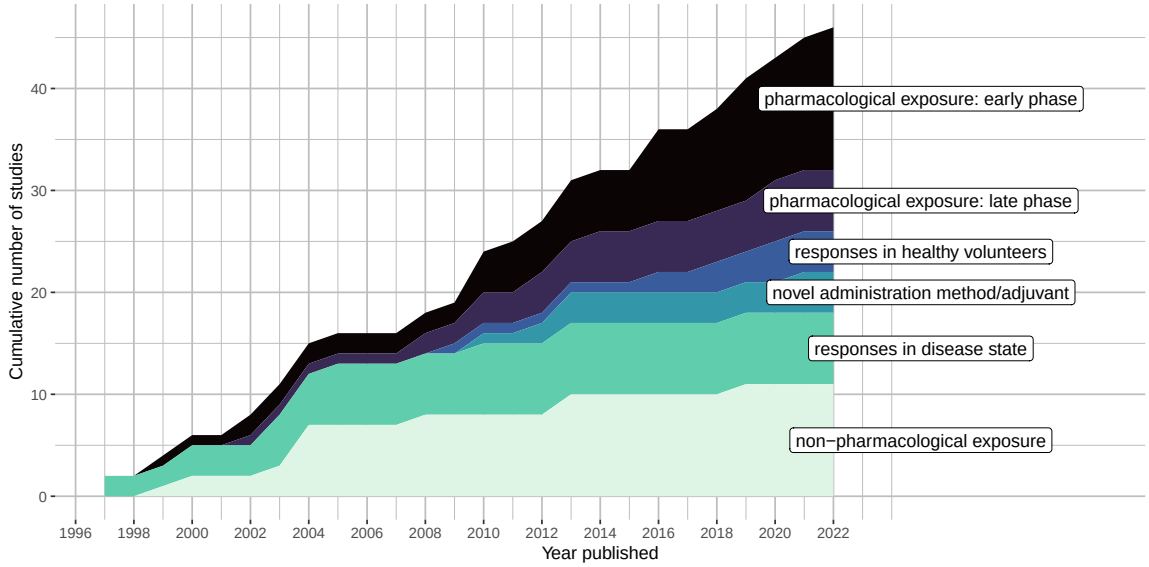

Figure 3.2: Experimental HIC studies employing KLH published between 1 January 1994 and 1 April 2022. x-axis: year published; y-axis, cumulative number of studies within each application category. Image reproduced from [38].

3.3.3 KLH challenge regimens

GMP-grade, sub-unit KLH was used in 6 of 14 studies (43%, all from a single supplier: Biosyn), which was administered with an adjuvant in three studies (aluminium hydroxide or Montanide ISA-51). HMW-KLH was used in three studies (21%), and an unspecified formulation in five studies (36%). The most common primary KLH dose used was 1000 µg (nine studies, range: 100 µg to 2500 µg), and was administered by the subcutaneous route in seven studies, intradermal route in two studies, intramuscular route in two studies, and via an unspecified route in two studies. Seven studies administered a single dose of KLH, while seven studies administered more than one dose. In two studies, intradermal KLH was administered as a rechallenge followed by DTH assessment (see outcome assessment, below). The primary dose of KLH was administered prior to the IMP in two studies, concurrently with initiation of the IMP in three studies, and following initiation of the IMP in seven studies. One study compared responses to KLH according to administration 14 days

before or 14 days after the initiation of the IMP (rituximab)[53], while another compared responses to KLH when administered concurrently to or 29 days following initiation of the IMP (an OX40 agonist).[56]

3.3.4 Outcome assessment and effect of study drug on KLH immune response

The response to KLH challenge was a primary outcome in two of the 14 (14%) studies identified, and a secondary or exploratory outcome in 12 studies (86%) which, consistent with their early phase status, generally focused on safety, PK, and efficacy biomarkers. 13 studies (93%) reported humoral responses to KLH, typically using ELISA with one study using a flow-cytometric bead-based assay.[59] Seven studies reported anti-KLH IgM responses in addition to IgG. Studies varied in the timepoints of antibody assessments and frequently assessed responses at multiple timepoints, most commonly 14 and 28 days (range 7-113 days) post KLH administration.

Of the two studies which elicited DTH by intradermal KLH rechallenge, both assessed responses clinically, via measurement of erythema, and/or induration using the 'ball-point pen technique' commonly-used for assessment of Mantoux/tuberculin skin tests.[97–99] In addition, Saghari et al. (2022) used non-invasive imaging techniques to objectively quantify the cutaneous DTH response (erythema and induration), and assessed induration, erythema, tenderness, and pain using a validated toxicity grading scale.[64, 94]

3.3.5 Other outcome assessments

All studies incorporated KLH challenge into a larger battery of pharmacodynamic assessments. Most studies reported drug target occupancy (e.g. on a relevant T-cell subset).[56–59, 61–64] Downstream PD outcomes included serum cytokine concentrations (commonly as a safety endpoint in CD28-targeting agents)[57, 58, 60, 62, 63], blood lymphocyte subsets [52, 56–60], viral reactivation (EBV and CMV)[57, 62], response to vaccination with other antigens (e.g. tetanus toxoid [53, 54, 64], pneumococcal polysaccharide vaccine,[60] bacteriophage ϕX174[52]), and effect of the IMP on pre-existing antibody titres (anti-tetanus toxoid IgG).[59] A smaller number of studies reported endpoints based on ex vivo assays, including cytokine release following antigen stimulation with pre-sensitised antigens (tetanus,[56] candida [55]) or non-specific stimulation (e.g. SEB, LPS, CD154).[57, 62, 97, 100] Studies incorporating patient populations reported disease specific-endpoints, including histological response (e.g. psoriatic plaques,[52] lymph nodes [97]), and disease activity endpoints assessed clinically or radiologically (e.g. autoimmune disease activity scores,[52, 60] tumour response,[56] thymus size,[55] lymph node size [97]).

3.3.6 Immunomodulatory mechanisms and proof-of-mechanism

Eight of the 14 IMP (57%) studied in early phase clinical trials were inhibitors of T-cell co-stimulation (Figure 3.3), including CD28 (FR104 and lulizumab)[57, 101], CD40 (dazodalibep)[61, 62], OX40L (amlitelimab/KY1005)[64], ICOSL (prezalumab)[59], dual ICOSL/CD28 (acazicolcept)[63], and CD80/86 (abatacept).[52] All but one of these studies demonstrated inhibition of anti-KLH antibody responses relative to placebo, with a non-statistically significant trend towards inhibition in the remaining study.[64] Exposure-

response relationships were assessed more-formally in some cases, e.g. by modelling antibody titres as a function of study drug exposure (AUC) or receptor occupancy—these assessments generally demonstrated clearer associations with the pharmacodynamic response to KLH, as compared to dose-response.[57, 59, 61, 64] Five studies of immunosuppressive IMP were performed in patient populations (phase 1b or 2), affording the potential for concurrent assessment of IMP modulation of both KLH response, and disease activity (or related biomarkers). Abrams et al. reported a study of abatacept (antagonist of the CD28-CD80/86 interaction) in patients with psoriasis.[52] In this study, suppression of total anti-KLH antibody titres (compared to a control group) was seen in one or more patients at all dose levels two weeks after secondary immunisation with 1000 µg HMW-KLH, although no clear dose-response was seen. Conversely, dose-dependent responses were seen in both clinical activity and histological markers of psoriatic activity, such as intralesional T cells. Sullivan et al. reported on the effects of AMG 557 (prezalumab), an ICOSL inhibitor, in an exploratory phase 1b study of SLE patients with inflammatory arthritis.[59] AMG 557 treatment resulted in a significant reduction of anti-KLH IgG responses following KLH challenge 1000 µg sub-unit KLH ID, 4 weeks apart). Notably, this effect was seen concurrently with acceptable target occupancy on circulating B cells (steady-state trough concentrations above IC99), but with no observed effect of SLE-related laboratory biomarkers (autoantibodies and complement concentrations) or disease activity scores. The authors asserted this lack of clinical response was consistent with the "mild, stable disease" status of the patients. St Clair et al. reported a phase 2 study of baminercept (a lymphotoxin β receptor fusion protein) in patients with Sjogren's syndrome challenged with sub-unit KLH 1000 µg with Montanide ISA-51.[60] There was no evidence of suppression of anti-KLH IgG responses in the intervention group, although mean anti-KLH IgM responses were higher c.f. control

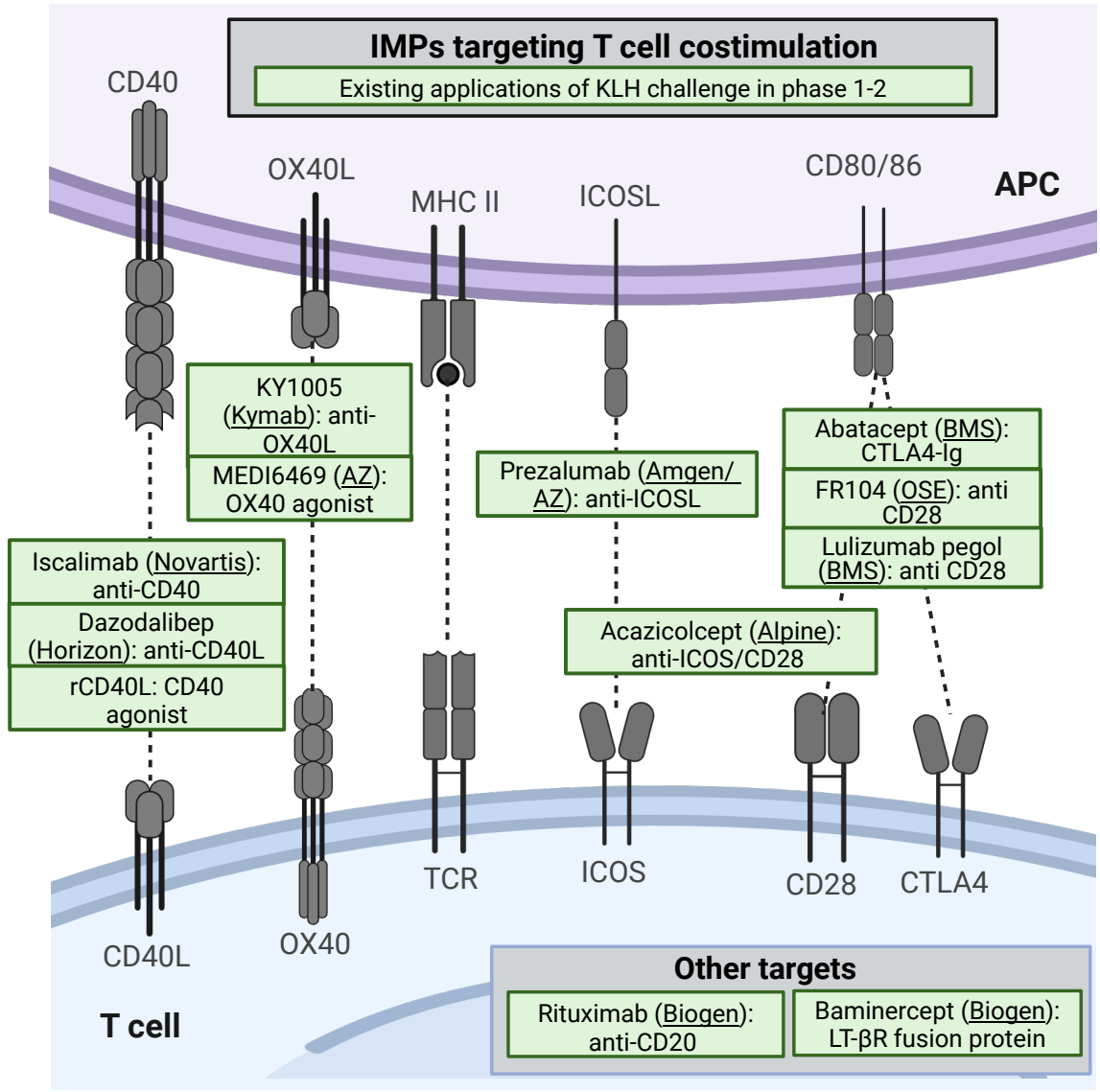

Figure 3.3: Molecular targets of early phase clinical trials of investigational medicinal products (IMP) applying KLH challenge 1994-2022. Most identified IMP target T cell co-stimulation (receptor interactions displayed graphically), top: antigen presenting cell, bottom: T cell). Created with BioRender.com. Image reproduced from [38].

(statistical significance not reported), postulated by the authors as possibly consistent with an effect on antibody class-switching. No effect on disease activity was observed. Two studies reported antibody responses to HMW-KLH in studies of the chimeric anti-CD20 mAb rituximab, although no clinical correlates could be derived from these studies, with the KLH challenge regimen failing to elicit a response in either control or intervention groups in one study of patients with relapsed low grade lymphoma,[53] and another study where the focus was on response to immunisations (e.g. for understanding infection risk) and where no clinical responses were measured.[54]

Three studies reported application of KLH challenge in the context of (potentially) immunostimulatory drugs. Curti et al. conducted a non-controlled phase 1 study of an investigational OX40 agonist in a heterogenous group of patients with metastatic solid tumours, and demonstrated a statistically significant increase in anti-KLH IgG in the treatment groups.[56] Regression of at least one metastatic lesion was reported in 12/30 (40% of) patients, although any potential correlation with response to KLH within individuals was not reported. Jain et al. described a small case series (n=3) in which children with X-linked hyper IgM syndrome (an immune deficiency disorder caused by mutations in CD40 ligand gene) receiving treatment with recombinant CD40L (rCD40L) successfully mounted a KLH DTH response following study drug initiation, while no antibody responses to KLH were detected. Notably, in these patients, the DTH response on subsequent intradermal rechallenge was absent after a drug-free period, but returned following rCD40L re-initiation.[97] Finally, Smith et al. reported the use of recombinant human growth hormone (rhGH, Somatotrophin) in patients with advanced human immunodeficiency virus (HIV) infection, postulating potential beneficial effects on CD4+ T cell count and immune function, including response to KLH challenge.[55] Overall, the effects of the intervention on the primary

outcome were modest. No differences in in vitro lymphproliferative responses to KLH were discerned in either the rhGH or control (antiretroviral therapy only) groups.

3.3.7 Concordance with preclinical immune challenge studies

10 of 14 studies (71%) compared observed IMP effects on anti-KLH antibody responses to those observed in animals receiving challenges with T-cell dependent antigens (including KLH) and the same IMP, or a species-specific analogue (table 3.3). In all cases, the response to the IMP was qualitatively similar between the human and animal studies.

Two studies used model-based approaches to bridge KLH challenge responses between species for the purposes of dose-selection and/or exposure-response analysis. Saghari et al. used pharmacokinetic considerations and allometric scaling to inform KY1005 dose selection in their (human) study, based on exposures observed to cause maximal suppression of anti KLH-IgG in monkeys.[64] In studies of a CD28 antagonist (lulizumab pegol), Shi et al. demonstrated that in both human and monkeys, maximal suppression of anti-KLH IgG response occurred with dose regimens where $> 80\%$ receptor occupancy was maintained for at least 2 weeks.[58, 102]

3.3.8 Reporting quality

Identified early phase trials were assessed against the TIDieR checklist (table 3.2). While the rationale for using KLH challenge was explained in most studies (13/14, 93%), the rationale for key aspects of the challenge regimen (e.g. dose, use of adjuvants) was generally not. Two studies (14%) explicitly referenced previous studies to explain regimen decisions, i.e. justified KLH regimen on the basis of precedent.[53, 64] One study timed KLH rechal-

lenge to coincide with expected loss of receptor occupancy of the study drug at different dose-levels.[62] Other key aspects of KLH challenge regimens were frequently not described. For example, four studies (29%) failed to describe the formulation and supplier of the KLH administered. Two studies (14%) did not describe the route of KLH administration, while 12 studies (86%) did not describe the body site of KLH administration. One study did not describe the dose of KLH administered. Five studies (36%) did not refer to the conduct of KLH challenge in the study abstract.

3.3.9 Late phase clinical trials, post-marketing surveillance, and observational studies of immunomodulatory drugs

Six of the 42 studies (14%) described application of the KLH challenge model to assessing the activity of drugs in late-phase clinical trials and observational studies. Collectively these studies attest to suppression of the humoral immune response to KLH by anti-CD20 agents (rituximab and ocrelizumab),[66, 70] S1PR modulation (fingolimod),[68] IMPDH inhibition (mycophenolic acid)[65], and lack of effect of an anti-α4-integrin mAb (natalizumab).[69] In one observational study of renal transplant patients receiving prednisolone and cyclosporine, the use of everolimus as a third agent was associated with retention of primary humoral immune response to KLH, in contrast to complete suppression of response seen with mycophenolic acid as a third agent.[67]

Two studies in this category reported assessment of DTH as an endpoint. In one study, healthy volunteers (n=72) were randomised to receive either placebo or one of two dose levels of fingolimod prior to immunization with 100 µg sub-unit KLH plus aluminium hydroxide adjuvant, once per week for 3 doses, followed by a 10 µg intradermal KLH dose to elicit a

DTH response. Notably, fewer than two participants in each of the placebo and intervention arms (n=24 each) developed a positive reaction (defined in this study as induration diameter $> 5mm$)—precluding the possibility of detecting an effect of the study drug. In the second, the effects of alternate immunosuppressive regimens used in renal transplant patients on DTH, proliferative responses of PBMC to KLH ex vivo, and immunohistochemical staining of punch skin biopsy at the site of intradermal KLH rechallenge were investigated. 'Vigorous' DTH responses were observed in healthy volunteers (n=10) receiving 100 µg intradermal skin challenges with HMW-KLH, 14 days following a primary intramuscular dose of 1000 µg of HMW-KLH, in contrast to 0 of 6 renal transplant patients treated for rejection with an anti-CD3 mAb, and 2 of 7 patients receiving standard immunosuppressive regimens (prednisolone, cyclosporine A +/- mycophenolate mofetil).[65] Conversely, there appeared to be limited differences between healthy volunteers and patients in terms of PBMC proliferative responses to ex vivo KLH stimulation. Immunohistochemistry identified a reduction of ICAM-1 in the setting of immunosuppression with prednisolone/cyclosporine A, while E-selectin, VCAM-1, and CD3 expression was not affected.

3.3.10 KLH challenge in other contexts

26 of the 46 studies (57%) described application of the KLH challenge model in other contexts, including the effects of non-pharmacological exposures or interventions (e.g. age, ultraviolet exposure, exercise, and psychosocial stress)[78, 79, 81, 82, 84, 89, 93], and the effect of disease states [71, 74, 77, 86, 103] on KLH response (table 3.5). Two studies reported novel assays for assessing KLH-specific immune responses.[39, 88] The majority (21/26, 81%) of the studies in this category were published prior to 2014.

Selected studies identified in this category are relevant for development of the KLH challenge paradigm in general. Saghari et al. (2020) sought to characterise the performance characteristics of a single dose level KLH challenge protocol, with a particular focus on objective imaging techniques to quantify the DTH response.[94] In this study, healthy male volunteers were administered sub-unit KLH (Biosyn, Carlsbad, CA) 100 µg (n=12) with aluminium hydroxide adjuvant or placebo (n=3) via intramuscular injection, followed by 1 µg intradermal injection 21 days later. The DTH response was assessed at 48h, using clinical assessment (induration, erythema, tenderness, and pain), and by a variety of imaging modalities, including multispectral imaging (erythema and oedema), colorimetry (erythema), and automated 2D photography (erythema). This study was notable for the low dose of sub-unit KLH used to elicit the DTH response, and associated excellent tolerance of the regimen. The DTH response was not apparent by visual inspection but could be identified using skin imaging. This regimen was subsequently applied to a phase 1 trial of an OX40L antagonist wherein modulation of the cutaneous DTH response by the IMP was demonstrated.[64]

Oyelaran et al. investigated antibody responses using a 107 component carbohydrate antigen microarray assay utilizing samples from an earlier study of healthy men (n=14, younger and older cohorts) receiving 100 µg sub-unit KLH and aluminium hydroxide (900 µg).[39, 79] Results from this study supported the hypothesis that the anti-KLH antibody response in participants was largely directed at carbohydrate epitopes (rather than peptides), with considerable inter-individual variability in the magnitude and breadth of responses to specific antigens, even amongst those with similar total anti-KLH IgG response. In addition, the investigators identified a subset (13 of 107) of pre-immune antibody responses, which were statistically significantly inversely correlated with the magnitude of the post-immune anti-

KLH IgG response (at a p=0.05 level without adjustment for multiple hypothesis testing), suggesting one potential method for accounting for some of the (marked) inter-individual variability in anti-KLH antibody responses as a function of baseline covariates.

3.3.11 Use of adjuvants (all study types)

10 of the 46 studies identified (22%) administered KLH with an adjuvant, namely aluminium hydroxide (seven studies), Montanide ISA-51 (two studies), and poly-ICLC +/- fimaporfin (one study). In all but one study these adjuvants were administered with sub-unit KLH (with the formulation not specified in the remaining study).

One study compared immune responses with and without adjuvants. Miller et al. described primary antibody and cellular responses (proliferation assays and anti-IFNγ ELISpot) to KLH in healthy volunteers (n=37) and a heterogenous group of patients with malignancies following chemotherapy (n=14) or haemopoietic stem cell transplant (n=19).[83] In this study, participants were administered three different KLH containing regimens SC: (i) HMW-KLH 1000 µg (Intracel, Rockville, MD), (ii) Sub-unit KLH 1000 µg or (iii) Sub-unit KLH 1000 µg with Montanide ISA-51 (Seppic Inc, Fairfield, NJ). The allocation of challenge regimen was non-randomised, and instead occurred due to unreliable supply of the HMW-KLH product initially used, necessitating a change in supplier. HMW-KLH induced a potent antibody response in healthy volunteers (n=17), whereas healthy volunteers (n=10) receiving the sub-unit product failed to mount a significant antibody or cellular response. Emulsification of the sub-unit KLH with Montanide ISA-51 adjuvant safely elicited a 100% antibody response rate in healthy volunteers (n=10). Patients (n=34), who received either HMW-KLH or sub-unit KLH with Montanide, demonstrated impaired responses (most pro-

nounced in patients post haematopoietic stem cell transplant). The comparative effect of KLH formulation and Montanide on qualitative immune responses (e.g. TH1:TH2 skew) was not directly evaluated.

We identified one study reporting the application of a novel adjuvant strategy for promoting CD8+ T cell responses, via application of photochemical internalisation (PCI)—a technique to induce the release of antigen from intracellular vesicles into the cytosol of antigen presenting cells in response to cutaneous laser illumination.[95] In a phase I dose-finding study, participants received two fortnightly intradermal doses of 100 µg sub-unit KLH (and human papilloma virus peptides) combined with a fixed dose (50 µg) of the TLR3 agonist poly-ICLC (Hiltonol, Oncovir, Washington DC) and ascending doses of the PCI agent fimaporfin. A control group received antigen and poly-ICLC alone. All participants developed anti-KLH IgG responses, with higher titres observed in participants receiving PCI with fimaporfin 12.5 µg compared to those in the control group. 50% of participants receiving sub-unit KLH with poly-ICLC alone developed positive IFN-γ ELISpot responses, compared to 40-100% of participants (depending on dose level) in groups receiving fimaporfin PCI plus poly-ICLC where there was no clear relationship observed between fimaporfin dose and ELISpot response.

3.4 Discussion

3.4.1 Principal findings

In this systematic review we identified a per annum increase in studies employing KLH-HIC to assess the activity of immunomodulatory IMP in early phase clinical trials. Whilst

KLH challenge regimens were predominantly employed to evaluate IMP targeting T-cell co-stimulatory pathways (immunosuppressive and immunostimulant), the paradigm was applied to the study of drugs exhibiting diverse immunomodulatory mechanisms. Most studies of immunomodulatory IMPs sought to demonstrate modulation of the anti-KLH IgG response as a primary KLH-specific outcome.

3.4.2 Deficiencies in existing applications of KLH challenge

In most cases, the rationale for a particular aspect of KLH challenge study design was not described or was based on precedent set by a previous study. We observed marked and unjustified variability in fundamental aspects of study design. For example, primary KLH challenge doses varied 50-fold (100 µg to 5000 µg, with max dose of 2500 µg in early phase studies), while KLH doses used for elicitation of DTH response ranged 100-fold between 1 µg to 100 µg. We identified no studies in the review period which explicitly sought to characterise KLH-dose response, or to define an 'optimal' dose. The consequence of poorly optimised challenge protocols is exemplified by studies in which the challenge regimen failed to elicit a response in an acceptable proportion of participants in any arm (including placebo), precluding an evaluation of the IMP using this endpoint,[53, 55, 68] and studies observing high rates of injection site reactions to KLH.[91] Heterogeneity was also observed with other aspects of design, including site and route of KLH administration, use of adjuvants, timing of KLH challenge relative to IMP administration, and both the nature and timing of outcome assessments. For example, of the two studies of immunosuppressive IMPs targeting the CD40-CD40L interaction, one introduced KLH 14 days before the first IMP (dazodalibep) dose, while the other introduced KLH three days after the first IMP (Iscalimab) dose.[61, 62] The timing of KLH challenge relative to IMP would be ex-

pected to present distinct pathways elicited for potential modulation although this was not explored. Overall, the available data suggest improved immunogenicity when sub-unit KLH is co-administered with an adjuvant, although we identified no randomised controlled studies which evaluated this question. Although differential qualitative effects on the resultant immune response may be expected by varying the choice of adjuvant, this has yet to be systematically explored, and variability in participant characteristics, KLH challenge regimens, and methods of outcome assessment limits the conclusions which can be drawn by comparing responses across existing studies. In some cases, the circumstances of a particular study will have driven rational decisions regarding KLH challenge regimen (e.g. high KLH doses used in the study of X-linked hyper IgM patients),[97] however in most cases variability in design was at best unexplained and at worst, arbitrary. Standardisation of protocols (where appropriate) is likely to improve the generalisability and external validity of findings and in turn necessitates further research that systematically explores the operating characteristics of KLH challenge in order to develop a protocol that is fit-for-purpose for, or can be tailored to, a broad range of applications. A well-characterised paradigm with established performance characteristics would inform sample size calculation and allow leverage of prior information (e.g. in Bayesian study designs). The value of this approach is likely to be high, especially in the setting of early phase trials of immunomodulatory agents that typically recruit small numbers of participants and which measure immunological endpoints, which are classically characterised by large inter-individual variability—a combination of design characteristics which risks indeterminate results, erroneous inferences, and flawed development decisions. Standardisation of procedures would additionally facilitate benchmarking between different IMPs targeting either the same or alternate mechanisms, enhancing insights into fundamental human immunology afforded by controlled perturbation. These

considerations align with those seen in preclinical KLH challenge studies, including those performed in non-human primates and rats, where the effects of key covariates (e.g. study site, gender, KLH formulation, use of adjuvants) has been quantified and used to inform power and sample size calculations.[42, 43] Importantly, studies frequently failed to report key aspects of KLH challenge regimen, including formulation and dose of KLH. The absence of this (basic) information severely limits the reproducibility of study findings, and echoes ubiquitous reporting deficiencies identified throughout the medical literature.[104, 105] Established guidelines exist for writing protocols and reporting of RCTs (SPIRIT[106] and CONSORT [107] respectively, with extensions for early phase dose-finding trials in progress [108]) and for the reporting of study interventions (TIDieR).[51] There is however no guideline specific to the reporting of experimental medicine approaches such as HIC, and none that we could identify as fit-for-purpose (without modification) for assessment of the studies in this review. As we have demonstrated, HIC can be performed as standalone research in healthy volunteer or patient populations, or integrated into studies with broader objectives, e.g. phase I clinical trials of IMPs. A reporting guideline specific to HIC (standalone or integrated into other studies, such as clinical trials), with sufficient detail to provide clear direction to authors, reviewers, and editors, may therefore be of great benefit for improving research reproducibility.

3.4.3 From immunotoxicity assessment to proof of mechanism

A key observation of this review is the evolving application of KLH HIC from immunotoxicity assessment (mirroring its role in preclinical TDAR-DTH studies [42]) to the assessment of PoM of immunomodulatory IMPs. In their systematic review of KLH HIC studies published between 1994-2013, Swaminathan et al. described applications which demonstrate

the role of KLH as a probe for gauging the general immunosuppressive effect (manifesting as anti-KLH IgG suppression) of various natural exposures, disease states, and assigned interventions (both pharmacological and non-pharmacological), but notably, identified no early phase clinical studies.[93] The increasing interest in use of KLH challenge to answer specific mechanistic questions highlights the limitations of study protocols as currently enacted. The choice of primary (KLH-specific) outcome exemplifies this: anti-KLH IgG is a useful and easily measured endpoint in immunotoxicity assessment, where the coordination of multiple processes of the innate and adaptive immune responses is required for successful IgG synthesis. Here, anti-KLH IgG may be considered a sensitive marker of immunotoxicity, in a context where specific mechanisms of inhibition of antibody synthesis are of less interest. In contrast, when applied to immunomodulatory IMPs, or indeed to elucidation of basic human immunology, a higher resolution picture afforded by modern immunological techniques is arguably required. The inconsistent link between anti-KLH IgG response and other pharmacodynamic markers discussed above illustrates the need to identify and characterise more mechanistically relevant KLH challenge endpoints which are fit-for-purpose for the PoM assessment of specific immunomodulatory IMPs. To provide just one example, while the study of Sullivan et al. asserted that evidence of anti-KLH IgG modulation by AMG 557 (prezalumab) "demonstrates a PD effect of AMG 557 in subjects with SLE consistent with the biology of the ICOS pathway and supports further studies of AMG 557 as a potential therapeutic for autoimmune diseases", this drug failed to meet primary endpoints in a subsequent phase II trial in SLE, and development has since been discontinued. PoM assessment has been demonstrated to bridge an important translational gap between preclinical studies and clinical trials for selected disease indications, thereby improving the probability of success. The principles of PoM therefore may be an indispens-

able paradigm for decision making in early phase drug development. Successful application of PoM principles however requires careful attention to mechanism relevant immune pathways to demonstrate as clearly as possible whether the IMP under investigation is behaving in vivo as intended. A more ambitious goal would be to identify and validate pathways upregulated by HIC, which map to pathophysiological processes upregulated in discrete autoimmune/autoinflammatory conditions, but this is a remote prospect in the absence of careful fundamental characterisation of KLH challenge's operating characteristics. The KLH challenge paradigm is likely to be significantly strengthened through the use of modern immunological techniques on tissues relevant to the biophase of the IMP of interest e.g. skin biopsy, skin blister fluid, and lymph node aspirates, collected at relevant timepoints (e.g. primary response, DTH response).[109, 110] There is also an opportunity to explore the effects of adjuvants on the polarisation of immune response to KLH, and to evaluate whether varying adjuvants can upregulate pathways relevant to a broader range of pharmacological mechanisms. These approaches and outcome assessments lend themselves to a quantitative systems pharmacology approach, which allows the development and evaluation of hypotheses relevant to the IMP of interest, and to more explicitly define the mechanisms which are being assessed in, for example, the clinical measurement of erythema or induration in a DTH response. Such approaches would also be of value in the qualitative and quantitative translation of preclinical KLH challenge studies to early phase clinical studies employing the same IMP (as demonstrated in a limited number of studies identified in this review), and may improve understanding of the translational relevance of KLH challenge performed in different non-human species.[58, 64] Table 3.6 summarises key outstanding questions and potential opportunities for KLH HIC that may be addressed in future studies. Based on the results of this review, it may be premature to make specific recommendations for standardi-

sation of experimental protocols, prior to further work on the fundamental characterisation and optimisation of the paradigm.

Table 3.6: Outstanding research questions for KLH-based human immune challenge studies

Domain	Research question and potential value
Design of challenge regimen	
	Q1. What is the optimal dose (primary and intradermal rechallenge) of KLH for HIC? *Potential value:* Standardisation of dosing between studies to improve benchmarking and generalisability of findings.
	Q2. How does response to intradermal KLH following initial immunisation vary over time, and upon multiple re-exposures? *Potential value:* Establish the feasibility of within-subject designs for proof-of-mechanism assessment and determine the sensitivity of responses to rechallenge timing.
	Q3. How might different adjuvants influence immune response polarisation to KLH, and can this be used to activate mechanism- and disease-relevant immune pathways? *Potential value:* Controlled elicitation of specific immune responses (e.g. T_H1, T_H2, T_H17) may increase the relevance of the model for mechanistic evaluation of immunomodulatory investigational medicinal products.
Outcome assessment	
	Q4. Can assessment of KLH-specific immune responses in relevant tissues (e.g. skin biopsy, blister fluid, lymph node aspirate) provide enhanced insight over peripheral blood sampling? *Potential value:* Characterisation of stromal immune responses may yield more mechanism-relevant information on investigational medicinal product activity.
	Q5. What are the determinants of the nature and magnitude of the immune response to KLH, can this response be predicted, and how does it vary between individuals? *Potential value:* Inform study design decisions including sample size calculations and prior specification for Bayesian statistical analyses.
	Q6. What is the time course of systemic and stromal immune responses following primary KLH immunisation and subsequent rechallenge? *Potential value:* Identify optimal sampling timepoints aligned to specific mechanistic hypotheses.
Maximising learning from KLH-based HIC studies	
	Q7. Can quantitative, model-based approaches be used to integrate existing knowledge and characterise response to KLH and its modulation by investigational medicinal products?

Domain	Research question and potential value (continued)
	Potential value: Model-based inference may mitigate limited statistical power in early-phase studies and support more robust mechanistic conclusions than simple frequentist analyses.

3.4.4 Strengths and limitations

In this study we present a comprehensive survey of the KLH HIC literature, and our focus on applications to drug development represents a distinct contribution when compared to earlier reviews.[16, 40] We used a comprehensive strategy and standardised approach to extract relevant study features, although there are some limitations. Given that KLH challenge is not usually the primary focus of a published study, its use is frequently not highlighted in study titles, abstracts, or keywords. We have attempted to identify all relevant studies using a comprehensive search strategy, including cross-referencing with clinical trials registries and searching of reference lists, although it is possible some relevant studies have not been identified, especially where KLH challenge is not specifically referenced in the title or abstract of the study. We chose to focus on the application of KLH HIC to drug development, and as such our analysis of KLH challenge applications in other settings (e.g. effects of non-pharmacological exposures, physiological states, and diseases states) was more limited. As demonstrated in figure 3.2, the incidence of these studies has reduced since 2014, and as such the previous review by Swaminathan et al. gives a suitably comprehensive account of these studies.[16] The deficiencies in KLH challenge identified in our review are likely to have similar implications on the validity, reproducibility and immunological insights afforded through these applications of the paradigm. Finally, this review focused on HIC utilising KLH only. Numerous antigens can be used to elicit a T-cell dependent immune response in humans, including tuberculin/purified protein derivative, tetanus, varicella,

candida, and others—a full survey of candidate antigens for HIC was beyond the scope of this review, and it is possible that an alternative antigen may be preferred for HIC in certain contexts. KLH has several attributes which favour its use for HIC, including xenogenic origin, such that most volunteers are immunologically naïve to relevant epitopes (although a small proportion of KLH-unexposed people will have cross-reactive immune responses), longstanding use (including in the preclinical setting), excellent safety record, and availability of a GMP-grade product. The value of KLH HIC should be evaluated in the context of complementary information that may be elicited by other pharmacological assays (e.g. target occupancy assays), which may be deployed in the evaluation of IMPs in early phase clinical trials. Ultimately PoM determination should be based on a total weight of high-quality evidence assessment.

3.5 Conclusion

KLH has an established role in immunotoxicology assessment with demonstrated predictive value in multiple non-human species. The studies identified in this systematic review attest to the value of human KLH challenge as a platform for interrogating a broad range of pharmacological mechanisms, translating preclinical data, and ultimately informing decision making in early phase clinical trials. They also highlight significant methodological and technical heterogeneity, and a historical reliance on simplistic, mechanism-independent biomarkers as PD endpoints. Explicit determination of the operating characteristics of the model, expansion of the range of immunological pathways elicited through alternate immunization protocols, acquisition of relevant biological samples and application of modern immunological techniques to KLH-driven responses in multiple tissue compartments, and clear comprehensive reporting of the results of these studies will maximise the latent poten-

tial of the paradigm. More broadly, this review highlights that methodological heterogeneity and failure to exploit the access to tissue-level mechanism-relevant endpoints afforded by KLH challenge has impaired the translational utility of this paradigm to-date.

In the context of the DPhil project, this review has confirmed the need and opportunity for further optimisation of KLH challenge paradigm, and has provided a clear set of priorities to be further evaluated in the planned KLH challenge studies, which enable the standardisation, characterisation and methodological development necessary to permit tailored, appropriately powered, mechanism-dependent study design to optimise drug development decisions.

Chapter 4

Quantitative refinement of the KLH Challenge model: adjuvant selection, dose-response modelling and time-course assessment

Chapter summary

Building on the systematic review of KLH challenge presented in the preceding chapter, this chapter reports two linked randomised, single-blind HIC studies designed to quantitatively optimise the KLH challenge model for proof-of-mechanism applications. This work addresses Objective 3 of the thesis:

To further characterise and optimise the KLH immune challenge model to advance its suitability for early-phase Proof-of-Mechanism studies in healthy volunteers:

a. *To determine the effect of co-administered adjuvants (Alhydrogel and Montanide ISA-51) on the magnitude and $CD4^+$ T_h1 / T_h2 polarisation of the resultant systemic primary immune response to KLH in healthy volunteers.*

b. *To quantify the effect of KLH rechallenge dose and identify an optimal dose for future PoM studies.*

c. *To characterise the temporal evolution of cutaneous immune responses to KLH*

rechallenge using multi-parameter flow cytometry, and thereby identify timepoints optimally suited to interrogate mechanistically distinct immune processes.

We demonstrate that adjuvant selection is a critical determinant of systemic immunogenicity, with Montanide ISA-51 eliciting substantially stronger and more consistent KLH-specific T cell responses than Alhydrogel, despite comparable humoral priming. Using Bayesian E_{max} modelling of the cutaneous recall response, we identify a 10 µg intradermal rechallenge dose as optimal, balancing response detectability, sensitivity to modulation, and tolerability. Finally, temporal profiling reveals that the canonical 48 h delayed-type hypersensitivity timepoint captures a mixed myeloid–lymphoid infiltrate, whereas later timepoints (Day 5 and Day 14) provide a more T-cell-dominant tissue milieu. Together, these findings advance a fit-for-purpose KLH challenge framework suitable for tissue-level PD in early-phase drug development.

4.1 Introduction

The systematic review reported in the preceding chapter confirmed the utility of KLH challenge as a promising method for the elicitation of PD biomarkers relevant to contemporary immunomodulatory drug development. However, despite extensive historical use, the review highlighted a critical lack of rational optimisation and standardisation in study protocols, as evidenced by the substantial observed heterogeneity in immunisation regimens and rechallenge doses. This variability hinders the benchmarking of immunomodulatory drugs and prevents the rigorous statistical design required for PoM studies. Furthermore, with endpoints largely limited to responses in blood (mainly serology), the range of immunomodulatory mechanisms which can be precisely and specifically assessed is unnecessarily limited.

In the previous chapter we synthesised these limitations into a set of outstanding ques-

tions required to develop the KLH challenge model into a fit-for-purpose drug development tool with broad applicability. These questions largely fell into two categories: (i) optimisation of the challenge regimen (specifically regarding adjuvant selection and antigen dose) and optimisation of outcome assessment (tissues, timepoints, and assays).

To address these questions we undertook a comprehensive characterisation of the KLH response through two randomised parallel-group HIC studies (KLH1 and KLH2). In KLH1, we focused on the magnitude and character of the response. While prior studies suggest adjuvants enhance immunogenicity, no study has systematically compared the effects of distinct adjuvants, such as the Th2 skewing aluminium hydroxide (Alhydrogel) versus the Th1/Th2 balanced water-in-oil emulsion Montanide-ISA 51, on the resulting systemic and cutaneous immune profile in humans.[83] Furthermore, while dose-dependency is assumed based on first principles and indirect comparisons between studies, the precise relationship between intradermal rechallenge dose and clinical response has not been modelled. Determining the shape of the dose-response profile is critical to identify a standard rechallenge dose which maximises consistency while avoiding saturation, thereby facilitating sensitive detection of pharmacological modulation.

In KLH2, we focused on the temporal evolution of the response. The commonly-selected 48 h clinical endpoint for delayed type hypersensitivity captures a specific moment in the dynamic immune response ('peak' macroscopic inflammation) but is insufficient to fully describe the evolution of specific immune processes relevant to many therapeutic targets, particularly for mechanisms targeting T cell differentiation, persistence, or tissue residency. By extending observation to day 5 and day 14, we sought to map the kinetics of the cellular infiltrate, determining optimal observation windows for key immune processes, such as myeloid and T cell infiltration, and local adaptive responses at the site of inflammation.

Accordingly, this chapter reports the results of these investigations, with the following specific objectives:

1. Compare the effect of two common adjuvants with distinct mechanisms of action (Montanide and Alhydrogel) on the humoral and systemic T cell response following primary KLH immunisation.

2. Model the dose-response relationship between intradermal KLH rechallenge and the peak (48 h) clinical response to determine the ED_{50} and identify an optimal dose for future studies.

3. Characterise the cellular composition of the cutaneous recall response, and its temporal evolution from 48 hours to 14 days post-challenge, including the comparative effects of primary adjuvant on this response.

4. Determine the agreement between established clinical assessment methods for DTH responses and objective non-invasive imaging modalities (LDI and multispectral imaging).

5. Characterise the participant experience to determine the overall acceptability and feasibility of the optimised challenge protocol.

4.2 Methods

4.2.1 Study design and participants

We performed two linked randomised, single-blind, parallel-group HIC (KLH1 and KLH2) to characterise the immune response to Keyhole Limpet Haemocyanin (KLH).

In the first optimisation study (KLH1), healthy volunteers ($n = 21$) were randomised (3:6:6:6) into a control group (Placebo) or one of three active intervention groups. The active groups received a primary intramuscular (IM) immunisation of 1000 μg subunit KLH (Immucothel, Biosyn) into the left deltoid, administered either alone (Unadjuvanted), or with one of two adjuvants: 900 μg Aluminium oxyhydroxide (Alhydrogel, Croda, Denmark; Th2-skewing) or Montanide ISA-51 (Seppic, Paris; 1:1 v:v; balanced Th1/Th2).[111] The control group received normal saline.

At Day 28, all participants in KLH1 received intradermal (ID) rechallenges on the volar aspect of the forearms bilaterally. To determine the dose-response relationship, six dose levels were administered: saline (negative control), 1, 3, 10, 30, and 100 μg of subunit KLH.

In the subsequent time-course study (KLH2), participants received an adjuvanted primary sensitisation regimen (1000 μg KLH + Montanide or KLH + Alhydrogel, 1:1 randomisation, n=12) followed by a standardised intradermal rechallenge dose (10 μg) to facilitate temporal profiling of the response, assessed at day 5 and day 14.

Inclusion and exclusion criteria are detailed in Table 4.1. All participants provided written informed consent prior to enrolment.

4.2.2 Randomisation, allocation, and masking

Participants were randomised using permuted block randomisation (block sizes of four and eight) generated using the `blockrand` package (version 1.5) in R. Allocations were capped to the maximum specified group size, allowing for additional recruitment in the event of dropout. Allocation cards were inserted into sequentially numbered, opaque, sealed envelopes and opened only following consent and confirmation of eligibility. Participants were masked to primary group allocation and the specific sequence of rechallenge doses admin-

Table 4.1: Inclusion and exclusion criteria for study participation

Inclusion criteria	Exclusion criteria
• Willing and able to give informed consent and comply with the study protocol. • Age 18–45 years (inclusive) at consent. • Healthy by medical history, examination, vitals and labs. • BMI 18–35 kg/m^2 (inclusive). • Females of childbearing potential: negative pregnancy tests and effective contraception until 60 days post–last dose. • Males with female partners of childbearing potential: effective contraception from first dose until 60 days post–last dose. • Sufficient English to complete consent procedures.	• Antibiotic/antiviral therapy for serious illness within 30 days. • SARS-CoV-2 infection within 30 days (PCR or lateral flow confirmed). • Immunosuppressant/immunomodulatory agents (systemic or topical) within 3 months. • Chronic conditions affecting immune responses (e.g., diabetes, significant atopy) or any condition interfering with study per investigator. • Tattoos, naevi, or skin abnormalities (e.g., keloids/history of keloids) that may interfere with assessments. • Pregnancy or breastfeeding. • Allergy to KLH, aluminium hydroxide, Montanide ISA–51, related adjuvants, or challenge components; shellfish allergy. • History of schistosomiasis (confirmed or highly suspected, or potential KLH cross–reactivity). • Previous exposure to Keyhole Limpet Haemocyanin. • Recreational sun–bathing/sun–bed use on the arm (wrist to shoulder) within 7 days of screening. • Needle or minor–procedure phobia. • Current smoker or using nicotine replacement therapy. • Vaccination within 2 months prior to screening or anticipated before end of follow–up. • Any other significant disease, disorder or finding that may put the participant at risk, affect participation, or impair data interpretation.

istered on the forearm.

4.2.3 Interventions and schedule of activities

An overview of the study schedules is shown in Figure 4.1 (KLH1) and Figure 4.3 (KLH2).

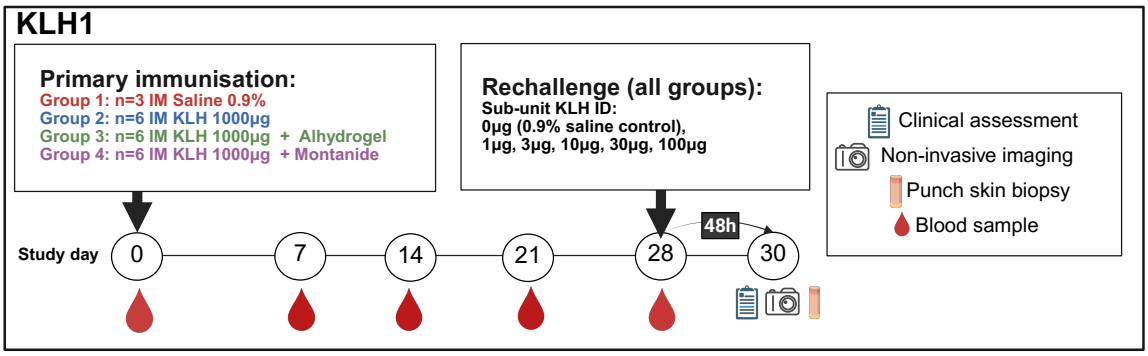

Figure 4.1: KLH1 study: interventions and schedule of activities. Schematic overview of primary immunisation, intradermal rechallenge, clinical assessments, and tissue sampling timepoints for the KLH1 cohort.

Primary immunisation and rechallenge On Day 0, participants received the assigned investigational medicinal product (KLH with/without adjuvant, or placebo) via deep intramuscular injection into the left deltoid muscle as per the KLH1 and KLH2 study schedules (figures 4.1 and 4.3).

Intradermal Rechallenge Participants in both studies received intradermal rechallenge as determined by the KLH1 and KLH2 study protocols.

- **KLH1 (Dose-ranging):** Participants received six intradermal injections on the volar forearms (3 per arm). These comprised five active doses of KLH (1, 3, 10, 30, 100 µg) and one negative control consisting of 0.9% saline (figures 4.1 and 4.2).
- **KLH2 (Time-course):** Participants received active rechallenge with 10 µg KLH on their volar forearms, on days 28 and 37, resulting in rechallenge lesions aged 5 days and 14 days at the time of assessment and biopsy on day 42. Note that for KLH2

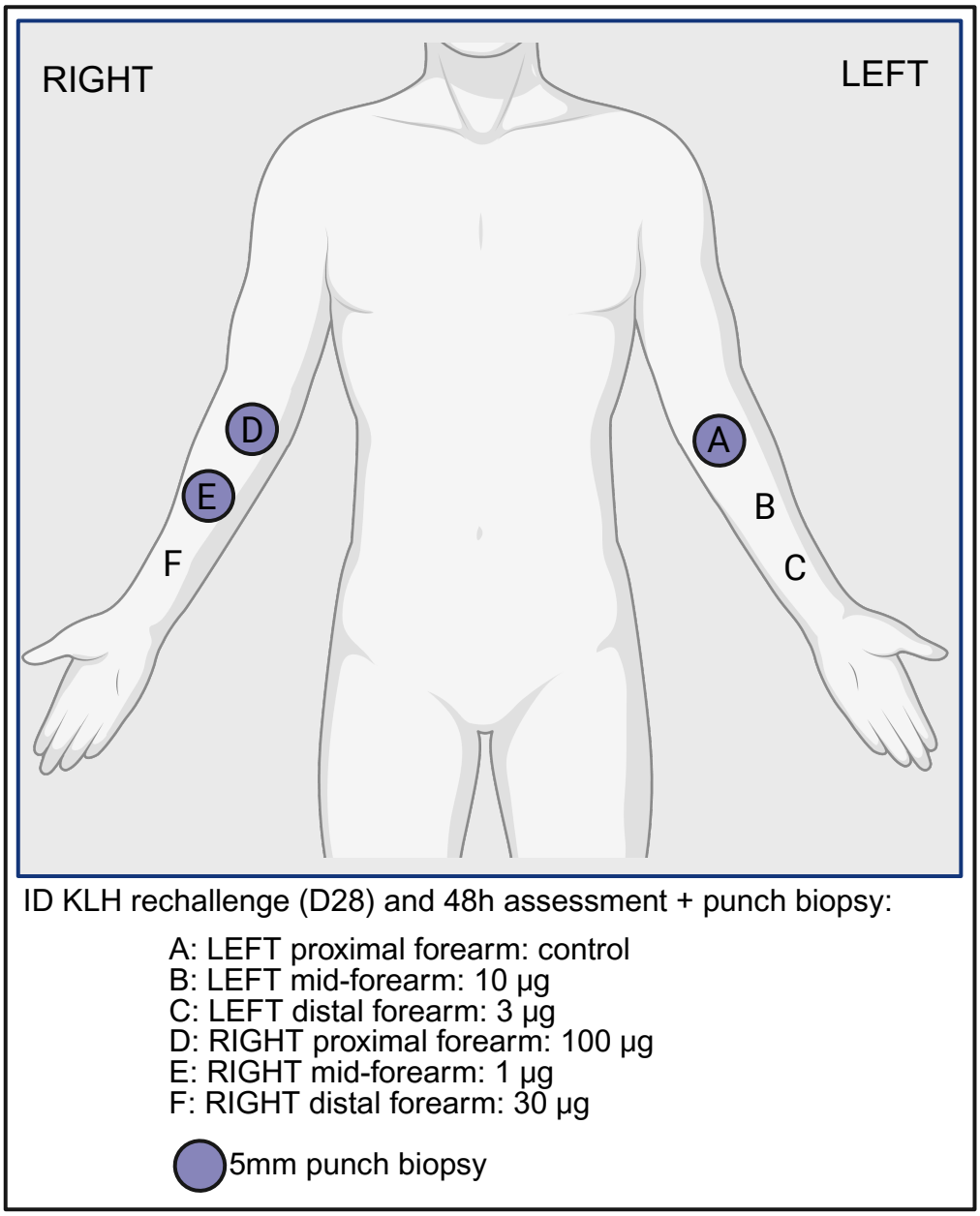

Figure 4.2: KLH1 study: intradermal rechallenge and biopsy sites. Participants received intradermal KLH rechallenge on Day 28, with doses ranging from 0 to 100 μg administered across predefined sites on the volar forearms.

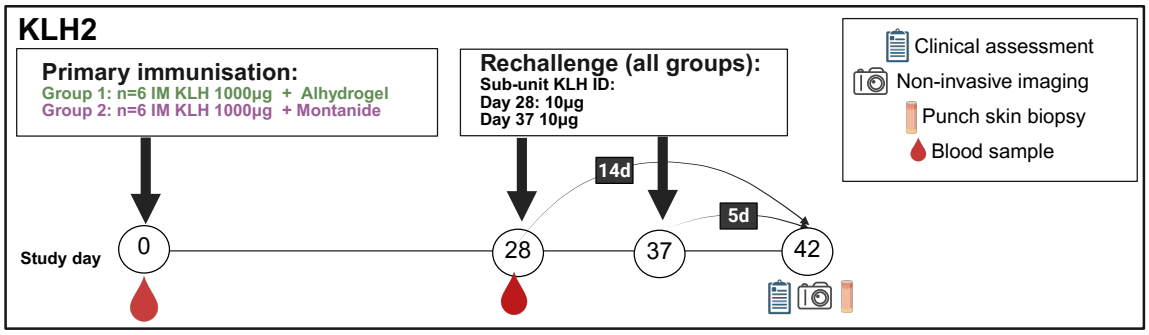

Figure 4.3: KLH2 study: interventions and schedule of activities. Schematic overview of primary immunisation, intradermal rechallenge, clinical assessments, and tissue sampling timepoints for the KLH2 time-course cohort.

the negative control site biopsied on day 42 was of unchallenged skin (figures 4.3 and 4.4).

4.2.4 Specimen collection and handling

Phlebotomy Peripheral blood samples were collected for analysis of serum (serum separator tube, BD), peripheral blood mononuclear cells (PBMCs, lithium heparin tube, BD), and RNA expression (Tempus Blood RNA tubes, ThermoFisher Scientific) at timepoints indicated in the study schedules. Blood samples were processed within 4 hours of collection. PBMCs were isolated by density centrifugation (Lymphoprep, STEMCELL). Serum and plasma samples were stored at -80 C, and PBMC samples were cryopreserved in 10% dimethylsulfoxide (DMSO) in fetal bovine serum (FBS) in liquid nitrogen until analysis. All laboratory assays were performed in the Fullerton Group laboratory according to standard SOPs.

Skin biopsy Punch biopsies were collected from challenge and control sites at 48 h post-rechallenge (KLH1 study) or five and 14 days post rechallenge (KLH2 study). For KLH1, 5mm punch biopsies were collected, which were bisected and processed in parallel for enzy-

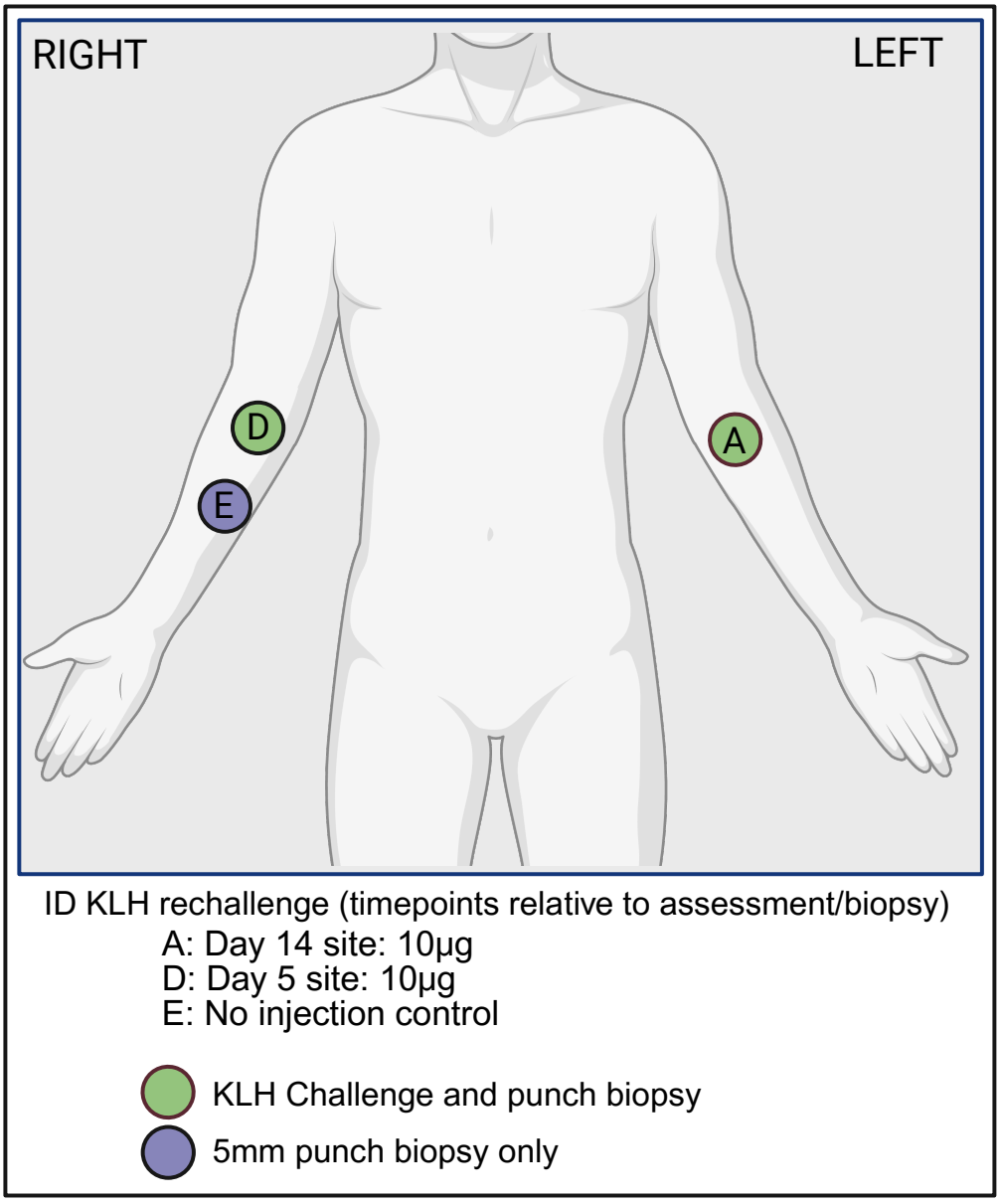

Figure 4.4: KLH2 study: intradermal rechallenge and biopsy sites. Participants received 10 µg intradermal KLH rechallenge on Days 28 and 37, with skin biopsies collected on Day 42, yielding samples from rechallenge lesions aged 5 and 14 days.

matic digest (flow cytometry, and bulk RNA sequencing in KLH1 study only) and histology (results not reported here). For KLH2, 6mm punch biopsies were collected and divided into four, with one quarter of the biopsy used for tissue digestion and flow cytometry. Digested cells were stored in liquid nitrogen until further use. Note that in both studies negative controls were taken (figures 4.2 and 4.4) however these differed between studies: for KLH1 the negative control was a site receiving 0.9% saline injection 48 h prior, while for KLH2 the negative control site was taken from uninjected skin.

4.2.5 Non-invasive assessments

4.2.5.1 Clinical assessment of cutaneous responses

Oedema and erythema The mean diameter of skin induration was measured using the ballpoint pen technique recommended for the assessment of Mantoux test reactions.[112] In this assessment a ballpoint pen is pressed lightly over the surface of the skin until a small amount of resistance is encountered. This point is then marked. The lesion is then approached and marked from the opposite side in the same manner, and the distance between the two points is measured with a ruler. The procedure is repeated at 90 degrees to this measurement and the mean of these measurements is calculated. A similar process was followed for the measurement of erythema, based on visual inspection of the extent of erythema relative to the surrounding skin.

4.2.5.2 Non-invasive imaging of cutaneous responses

We assessed the potential adjunctive value of two objective non-invasive imaging modalities to quantify the magnitude and extent of cutaneous inflammation.

Moor laser doppler imaging (LDI) Cutaneous blood flow was quantified using a MoorLDI2-BI imager (Moor Instruments, Axminster, UK). Image analysis was performed using the moorLDI Laser Doppler Review software (V6.2). For each image, a background region of interest (ROI) of $> 100\text{mm}^2$ was selected from adjacent, non-challenged skin. The mean flux of this background ROI was calculated and subtracted from the total image flux to generate background-corrected perfusion images. Following background subtraction, the specific reaction site was defined by manually tracing the visible lesion using the software's polygonal ROI tool. The primary outcome measure was the background-corrected mean flux within this specific ROI. Lesion dimensions were verified using the software's linear measurement tool.

Miravex Antera multispectral imaging The Antera 3D camera (Miravex, Dublin, Ireland) utilises multi-directional illumination and multispectral analysis to quantify numerous metrics related to chromophore distributions in the skin. For this analysis we focused on the quantification of redness via CIELab a* parameter. A broad circular ROI of 5 cm diameter was centered on the rechallenge site and the CIELab a* value was calculated using the system analysis software.

4.2.6 Participant experience

At the completion of the KLH1 study participants were administered a survey which assessed motivations for participating, participant experience, and general experience of the study.

4.2.7 Laboratory assays

Anti-KLH ELISA Anti-KLH IgG responses were assessed on cryopreserved serum using a validated commercial ELISA kit (Human Anti-KLH IgG ELISA, Alpha Diagnostic

International, San Antonio, TX, USA) according to the manufacturer's instructions. For the assay, samples were initially diluted in Working Sample Diluent to stabilise antibody activity, followed by a further dilution (1:200 or greater, as required) in Low NSB Sample Diluent to minimise non-specific binding. Diluted samples, along with calibrators (10-100 U/mL) and controls, were added to microwells pre-coated with sub-unit KLH (Immucothel, Biosyn) and incubated for 60 minutes at room temperature. After four wash cycles to remove unbound components, a horseradish peroxidase (HRP)-conjugated anti-human IgG detection antibody was added and incubated for 30 minutes. Following a final wash step (five cycles), a chromogenic substrate (TMB) was added, and the plate was incubated in the dark for 15 minutes. The reaction was terminated with Stop Solution, and absorbance was measured at 450 nm using a microplate reader. Antibody activity (U/mL) was calculated by interpolation from a calibrator curve generated using the kit's standards. While the primary outcome analysis utilised continuous antibody titres, a binary "seropositive" status was also defined for indicative purposes as a value exceeding the mean plus three standard deviations (3 s.d.) of the baseline (pre-immunisation) response

For the purposes of this study a positivity threshold was considered a value 3.s.d. above the mean log_{10} transformed response. Distributions of responses at baseline, in the control group, and following vaccination were used to confirm that this method yielded clear separation between non-immune and immune samples.

KLH ELISpot (IFN-γ / IL-4) Antigen-specific cytokine secretion was assessed using human IFN-γ and IL-4 ELISpot kits (Mabtech AB, Nacka Strand, Sweden) according to the standard operating procedure. Cryopreserved PBMCs were thawed and rested for 2 hours in complete medium (RPMI 1640 supplemented with 10% fetal bovine serum and

1% Penicillin-Streptomycin) at 37°C with 5% CO_2 prior to plating. Cells were seeded in triplicate at 3×10^5 cells per well in plates pre-coated with capture antibodies specific for the target cytokine. Cells were stimulated with subunit KLH (Immucothel, Biosyn, Fellbach, Germany) at a final concentration of 100 μg/mL. Anti-human CD3 mAb was used as a positive control to confirm cell viability and functionality, while media alone served as the negative control to establish background reactivity. Plates were incubated for 18–20 hours at 37°C in 5% CO_2 to allow for immediate capture of the secreted target protein. Following incubation, captured analyte was detected using a detection antibody conjugated to an enzyme, followed by a substrate that produces a precipitating product to visualise spots. Spots were counted using an automated ELISpot reader (AID, Strassberg, Germany). As per the laboratory SOP, results were calculated by subtracting the mean background (media only) response from the mean of the antigen-stimulated wells and normalised to report Spot Forming Cells (SFC) per 10^6 PBMCs. While the primary analysis treated the ELISpot response as a continuous quantitative variable, a binary classification of response status was also determined for indicative purposes; a positive response was empirically defined as a value exceeding the mean plus three standard deviations (3 s.d.) of the baseline (pre-immunisation) responses.

Flow cytometry (skin) Single-cell suspensions derived from skin biopsies were first stained for viability using the Zombie NIR™ Fixable Viability Kit (BioLegend) and incubated with a panel of fluorochrome-conjugated antibodies targeting surface antigens for 30 minutes at 4°C in the dark. Following surface staining, cells were washed twice with FACS wash buffer. To enable assessment of nuclear and cytoplasmic targets, cells were fixed and permeabilised using the eBioscience™ Foxp3 / Transcription Factor Staining Buffer Set (In-

vitrogen; 00-5523-00) according to the manufacturer's instructions. Intracellular staining was subsequently performed for 30 minutes at 4°C in the dark.

Data acquisition was performed on a Sony ID7000 spectral flow cytometer (Sony Biotechnology) running ID7000 software (v2.02). High-dimensional spectral data were unmixed and analysed using FlowJo software (version 10, FlowJo LLC). A complete list of antibodies, clones, and fluorochromes used in this panel is provided in Table 4.4.

4.2.8 Sample size justification

We identified eight studies administering a 1000 µg dose of subunit KLH as primary immunisation; two (Kaufman et al., Miller et al.) reported extractable binary serology.([69, 83]. In Kaufman (placebo cohort, fortnightly 3 times 1000 µg KLH), 9/24 (38%) were seropositive at Day 14; we therefore assumed a lower rate after a single 1000 μg dose. In Miller (non-randomised comparison), 0/10 responded at Day 28 after 1000 μg unadjuvanted KLH, versus 10/10 after 1000 μg with Montanide ISA-51. We identified six studies with alum-adjuvanted subunit KLH; none used 1000 μg: four used 100 μg (single dose) and two used three-dose series (weekly/fortnightly). Boulton et al. (n=24 controls) reported 96% achieving $\geq$2-fold anti-KLH IgG at Day 28.[68] On this basis, we considered $\leq$16.7% plausible for unadjuvanted subunit KLH and $\geq$83% for adjuvanted KLH. For the primary outcome in Phase 1, n=6 per arm provides 80% power to detect $\geq$82% response in adjuvanted groups vs 16% in unadjuvanted (one-sided α=0.05, chi-squared).

4.2.9 Statistical analysis

Statistical analyses were performed using R (version 4.4.0) in RStudio (2024.04.1+748). For frequentist hypothesis tests a p value of $<$0.05 was used to assess statistical significance. The

cohort and resultant immune responses were described using descriptive and comparative statistics.

4.2.9.1 Serological and T cell response.

To capture the total magnitude of the immune response over the observation period, the Area Under the Curve (AUC) for the interval Day 0 to Day 28 (AUC_{0-28}) was calculated for each participant using the trapezoidal rule. For serological analysis, antibody titres were log_{10}-transformed prior to AUC calculation to satisfy normality assumptions (units: log_{10} titre $\cdot$ days); ELISpot responses were analysed as raw spot forming cells (SFC) integrated over time (units: SFC $\cdot$ days).To determine the effect of adjuvant selection on immunogenicity, we employed two complementary modelling approaches. First, pairwise differences between the active study arms (Unadjuvanted KLH, KLH + Alhydrogel, KLH + Montanide) were assessed using a one-way ANOVA followed by planned contrasts with Holm-Bonferroni correction for multiple comparisons. Second, to rigorously compare the two adjuvanted regimens while accounting for biological variation, we performed an Analysis of Covariance (ANCOVA). This model compared the AUC_{0-28} between the Montanide and Alhydrogel groups, adjusting for age, sex, and baseline CMV serostatus using Type II Sums of Squares.

4.2.9.2 Intradermal KLH Dose-response modelling

Modelling approach We modelled the association between KLH rechallenge dose and induration at 48 h using a Bayesian hierarchical Emax framework (where Emax is the maximum response, and ED50 is the dose corresponding to half-maximal response) to account for repeated measures within participants and between-group differences across primary

sensitisation regimens. Analyses focused on the dose-response amongst individuals with phenotypic evidence of anti-KLH immunity, defined as an induration diameter of ≥5 mm at ≥1 non-zero rechallenge doses. The control responses were excluded from the data based on the prior assumption that they would not lead to an induration response; we instead fixed the model baseline response $E_0 = 0$. Models were fit to the dose scaled to the interval (0,1] to improve sampling efficiency, and then back-transformed to the natural scale for model summarisation.

Model specification For observation i from subject j in group g, with scaled dose $d_i^* \in (0, 1]$, the mean curve is:

$$\mu_i \;=\; \mathrm{Emax}_j \, \frac{d_i^*}{\mathrm{ED50}_g^* + d_i^*}.$$

Subject-level multiplicative heterogeneity on the maximal response (Emax) is defined as:

$$\mathrm{Emax}_j \;=\; \mathrm{Emax}_g \, \exp(\eta_j), \qquad \eta_j \sim \mathcal{N}(0, \sigma_\eta).$$

Group-level hierarchies The group-level parameters are modelled on the log scale using non-centred parameterisation:

$$\log \mathrm{Emax}_g = \log \mathrm{Emax}_{\mathrm{pop}} + \tau_{\mathrm{Emax}} \cdot z_g^{(\mathrm{Emax})},$$

$$\log \mathrm{ED50}_g^* = \log \mathrm{ED50}_{\mathrm{pop}}^* + \tau_{\mathrm{ED50}^*} \cdot z_g^{(\mathrm{ED})},$$

where $z_g^{(\mathrm{Emax})}$ and $z_g^{(\mathrm{ED})}$ are independent standard normal deviates:

$$z_g^{(\mathrm{Emax})},\, z_g^{(\mathrm{ED})} \sim \mathcal{N}(0, 1).$$

Observation model The observed induration y_i is distributed log-normally around the mean curve:

$$y_i \sim \text{LogNormal}\big(\log \mu_i,\ \sigma_{\log}\big).$$

Priors Vaguely informative priors were selected based on a review of prior data and first-principles considerations (e.g., anatomical limitations of the maximum extent of a DTH response on the forearm).

$$\log \text{Emax}_{\text{pop}} \sim \mathcal{N}(\log(75),\ 0.20),$$

$$\tau_{\text{Emax}} \sim \mathcal{HN}(0,\ 0.40),$$

$$\sigma_\eta \sim \mathcal{HN}(0,\ 0.30),$$

$$\log \text{ED50}^*_{\text{pop}} \sim \mathcal{N}(\log(0.2),\ 0.60),$$

$$\tau_{\text{ED50}^*} \sim \mathcal{HN}(0,\ 0.35),$$

$$\sigma_{\log} \sim \mathcal{HN}(0,\ 0.50),$$

where $\mathcal{HN}(0, \sigma)$ denotes a half-normal distribution with scale σ.

Reported quantities We reported the natural-scale parameters Emax_{pop}, ED50_{pop}, and ED50_g (in µg), as well as interpretable dispersion summaries (e.g., observation CV% derived from $\sigma_{\log}$ and between-group SD% derived from the τ terms).

Posterior inference Models were implemented in Stan via `cmdstanr` using the Hamiltonian Monte Carlo (HMC) algorithm for posterior inference. We used four chains (1500 warm-up, 2500 sampling), adapt_delta = 0.99 and max_treedepth = 12. Initial values were set near prior centres to improve sampler efficiency. Diagnostic and summary plots leveraged

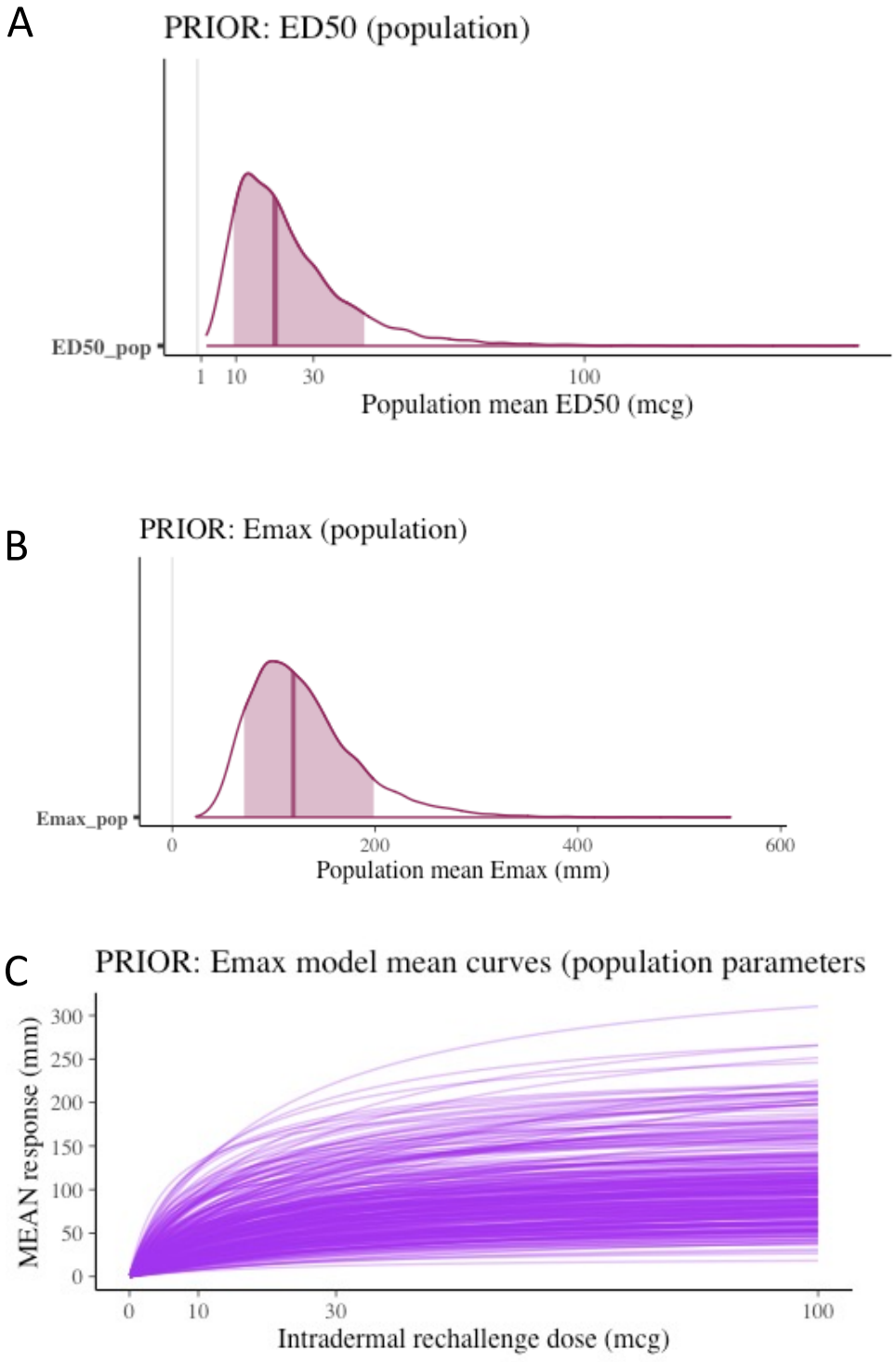

Figure 4.5: Prior distributions for population-level parameters in the Bayesian E_{max} dose–response model. (A) Prior density for the population maximal response (E_{max}); (B) prior density for the population half-maximal effective dose (ED_{50}); (C) prior predictive distribution for the mean induration response. The prior specifications allow substantial data-driven posterior updating while constraining responses to physiologically plausible ranges.

the R packages `Posterior` and `Bayesplot`.

Sampler diagnostics Sampler diagnostics were performed using in-built cmdstanr functions for divergent transitions, E-BFMI, rank-normalised split effective sample size and rank-normalised split R hat values as recommended. In addition we assessed chain trace and density plots, and pairs plots of key parameters to assess for pathological geometry.

Model checks Posterior predictive checks based on replicated draws (y_{rep}) were examined to assess agreement between observed data and the fitted model.

Model summaries For between-group inference, we summarised posteriors for Emax_g and ED50_g and computed comparisons (e.g., $Pr[\text{Emax}_{\text{montanide}} > \text{Emax}_{\text{alhydrogel}}]$) by comparing draws. To inform rechallenge dose selection for future studies, we performed posterior predictive simulations at candidate doses (1, 10, 100 µg) and evaluated the probability of an optimal response. An optimal response was defined as one which maximises the probability of a consistently detectable response (amenable to clinical non-invasive assessment, imaging assessment, and invasive assessment via skin biopsy) while avoiding excessively large responses which may be associated with adverse effects and which may be less sensitive to modulation. An indicative optimal response window was defined post-hoc as induration >10 mm and <60 mm.

4.2.9.3 Compositional analysis of skin biopsy cell populations

Sum-constrained beta–binomial models (R package `sccomp`) were used to analyse flow cytometry data from 48 h punch biopsies and assess effects of primary adjuvant, induration, and rechallenge dose. A full model description is given in the original publication. [113] Models assessed differential composition and variability while accounting for compositional

constraints, with HMC sampling in Stan (four chains; 1000 post–warm-up iterations). Convergence was assessed via $\hat{R}$<1.01 and effective sample size > 400; trace and density plots were inspected. Parameter estimates on the unconstrained (logit) scale were mapped to fold-change for interpretation; A 5% False Discovery Rate (FDR) was used for post hoc significance testing.

We developed a series of models:

Model set.

- **Model 1:** Induration (binary; ≥5 mm) → composition of major skin populations.
- **Model 2:** Induration (continuous) → composition of major populations (indurated sites only).
- **Model 3:** Primary adjuvant (Montanide vs Alhydrogel) and dose (1 vs 100 μg) → composition of major populations.
- **Model 4a:** Adjuvant and dose → composition of classical $CD4^+$ phenotypes.
- **Model 4b:** Adjuvant and dose → $CD4^+$ TF subsets (T-bet, GATA3, RORγt).

4.2.9.4 Performance of non-invasive imaging

Agreement with clinical assessment We compared the performance of non-invasive imaging metrics derived from the LDI and Antera systems to the mean diameter of induration, which served as the clinical reference standard. To assess the strength of the monotonic relationship between these distinct modalities, pairwise associations were quantified using Spearman's rank correlation coefficients (ρ).

Association with cellular readouts To determine the biological validity of non-invasive metrics, we assessed their association with the underlying cellular infiltrate at active rechallenge sites where paired skin biopsies were obtained (1 and 100 μg doses). We examined the correlation of each clinical and imaging metric with two key cellular outcomes: a) percent of $CD4^+$ T cells and b) percent of total monocytes. The predictive performance of clinical induration versus imaging surrogates was benchmarked via Spearman's ρ, allowing for the identification of the non-invasive metric that best reflected the true cellular magnitude of the response.

4.2.9.5 Transcriptomic analysis

RNA extraction, library preparation and sequencing Although the main focus of the skin analysis was evaluation by flow cytometry, we performed additional bulk RNA-seq analysis of samples yielding high cell numbers following digest to assess the potential for multi-modal assessment of skin responses via a single 5 mm punch biopsy. Digested samples yielding $\geq$400,000 viable cells were selected for bulk RNA-seq (100,000 cell aliquot). Total RNA was extracted using the RNeasy Micro Kit (Qiagen). Cryopreserved skin digest samples were processed according to manufacturer instructions. Purified RNA was sent to Azenta Life Sciences (Leipzig, Germany) for library preparation and sequencing. Poly(A)-selected libraries were generated and sequenced in 150 bp paired-end mode on a NovaSeq 6000 instrument (Illumina, San Diego, USA).

Alignment and transcript quantification Data were processed using nf-core/rnaseq (v3.14.0; doi: 10.5281/zenodo.1400710), part of the nf-core collection of workflows (Ewels et al., 2020), using reproducible software environments from the Bioconda (Grüning et al., 2018) and BioContainers (da Veiga Leprevost et al., 2017) projects. Raw reads were

trimmed to remove adapter sequences and low-quality bases with Trim Galore (v0.6.7), then aligned to the GRCh38.p14 human reference genome (RefSeq accession GCF_000001405.40) using STAR (v2.7.10a). Alignments were projected to the transcriptome and quantified with Salmon (v1.9.0), and transcript-level estimates were collapsed to gene-level counts using tximeta (v1.12.0). Across all samples, the mean alignment rate was 76.4% (range 70.7–79.2%), and the mean sequencing depth was 33.8 million (range 27.9–48.1 million) paired-end reads per sample.

Differential expression analysis Gene-level counts from Salmon were imported into R and linked to sample metadata. Due to the small number of highly cellular samples yielded from placebo and unadjuvanted KLH groups, and control and 1 μg rechallenge specimens, we restricted our analysis dataset to samples from sites receiving 100 μg rechallenges of participants in the adjuvanted KLH groups. Technical QC included inspection of library sizes, principle component analysis of voom logCPM values, and sample-sample correlation heatmaps. Lowly expressed genes were removed using `filterByExpr` (edgeR), and library sizes were normalised with the trimmed mean of M values (TMM) method. The primary comparison of interest was differential expression between KLH + Alhydrogel and KLH + Montanide groups. Differential expression was assessed using limma-voom with quality weights: voom-transformed $\log_2$ counts per million were fitted with a linear model including the primary adjuvant group, followed by empirical Bayes moderation with robust and trend options enabled. For the Montanide vs Alhydrogel contrast, $\log_2$ fold-changes and Benjamini-Hochberg (FDRs) were reported. Genes with FDR ≤ 0.10 and absolute $\log_2$ fold changes ≥ 1 were considered differentially expressed.

Geneset Enrichment Analysis using Singscore Due to the small number of samples available for RNA-seq analysis, absence of expression data from control sites, and absence of differential expression between primary adjuvanted groups (explained in Results, below), we sought to understand the landscape of immune activation elicited by intradermal rechallenge in sensitised individuals in general. Pathway-level activity was assessed on a single-sample basis using singscore. Immune-related gene sets were obtained from MSigDB via `msigdbr`, including Hallmark Immune Signatures (H collection), Reactome immune pathways (C2:CP:REACTOME), KEGG immune pathways (C2:CP:KEGG), and transcription factor target sets from the C3:TFT_LEGACY collection. A full list of the examined gene sets is given in the supplementary material (Table 4.24). Gene symbols were intersected with the voom expression matrix, and gene sets with 10–5000 expressed genes were retained. For each gene set, singscore was applied to rank-based expression values to obtain per-sample scores. Empirical p-values were derived by comparing observed scores to null distributions generated from 2,000 random gene sets of matched size per pathway.

4.2.10 Ethics Statement

The study protocol was approved by the UK NHS Research Ethics Committee (East of England - Cambridgeshire and Hertfordshire REC ref 22/EE/0150). All participants provided written informed consent before any study procedures.

4.2.11 Role of the funding source

The funder had no role in study design, data analysis, data interpretation, and writing of the report.

4.3 Results

4.3.1 Participants

For KLH1, a total of 24 participants were enrolled between March 2023 and April 2024, including three additional participants above the initial planned sample size to offset PBMC sample loss early in the study, which was traced back to a faulty serum batch used for the first five PBMC cryopreservations. All participants completed follow-up to Day 30 (Figure 4.6). The median age was 28 years (IQR 24.5-33.5), and 13/24 participants (54%) were female. CMV serology was measured in 22/24 participants, with 10/22 (45%) seropositive (Table 4.2). There were no material differences in age, BMI, sex, ethnicity, Fitzpatrick skin type, CMV serostatus, or baseline blood parameters between the groups (Table 4.6).

For KLH2, 12 participants were enrolled to facilitate time-course assessment of the cutaneous response. The cohort had a median age of 23.5 years (IQR 21.5–27.0) and was predominantly male (8/12, 67%). In contrast to the KLH1 cohort which was predominantly White (16/24, 67%), with skin phototypes ranging from I–IV, the KLH2 cohort was comprised of participants from Asian (4/12, 33%), Black (3/12, 25%), and mixed/other (5/12, 42%) groups, and included participants with higher skin phototypes (Fitzpatrick V–VI: 2/12, 17%) (Table 4.3).

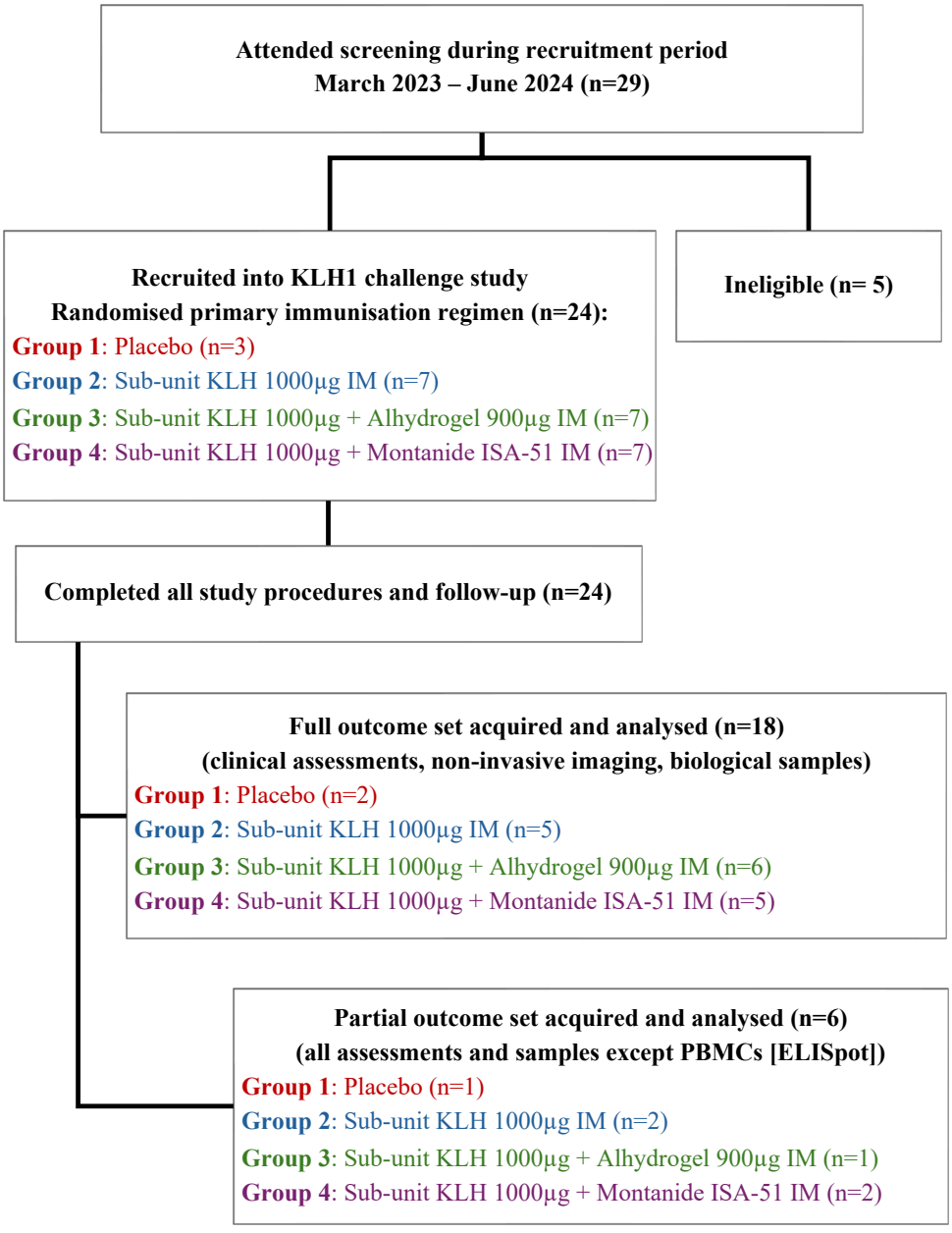

Figure 4.6: CONSORT diagram for the KLH1 study. Overview of participant enrolment, randomisation, follow-up, and timepoints of peripheral blood sampling according to challenge regimen.

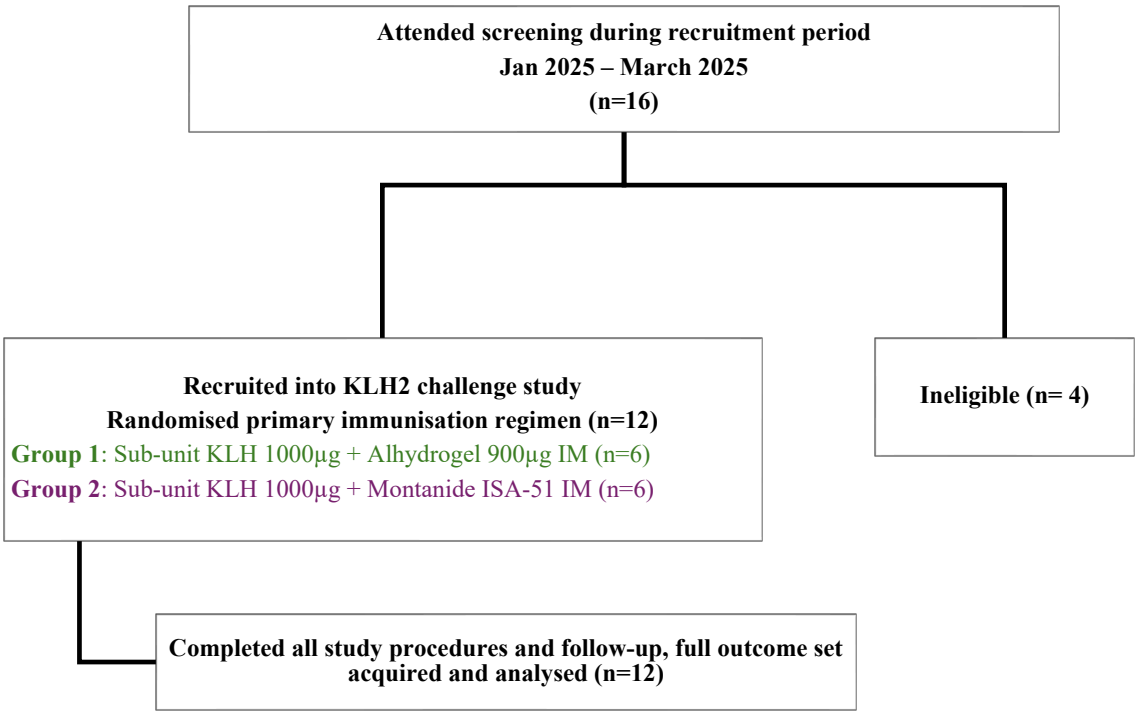

Figure 4.7: CONSORT diagram for the KLH2 study. Overview of participant enrolment, allocation to primary immunisation regimens, and follow-up through intradermal rechallenge and tissue sampling.

Table 4.2: KLH1: Demographics and CMV serostatus by randomisation group

Characteristic	N	Randomisation group				
		Placebo N = 3	**KLH 1000** N = 7	**KLH 1000 + Alhydrogel** N = 7	**KLH 1000 + Montanide** N = 7	**Overall** N = 24
Age (years)	24	28.0 [24.0, 38.0]	28.0 [23.0, 32.0]	28.0 [23.0, 34.0]	29.0 [27.0, 34.0]	28.0 [24.5, 33.5]
Height (cm)	24	182.0 [160.4, 189.0]	167.5 [162.0, 183.5]	169.0 [157.5, 177.0]	172.0 [168.5, 180.0]	171.0 [164.2, 181.0]
Weight (kg)	24	76.9 [61.6, 85.5]	70.4 [56.8, 89.6]	68.9 [61.4, 80.0]	77.6 [68.0, 87.2]	74.6 [66.2, 86.4]
BMI (kg/m^2)	24	23.9 [18.6, 29.9]	25.7 [20.7, 26.6]	25.5 [23.2, 28.0]	28.2 [22.9, 28.8]	25.6 [23.1, 28.4]
Sex	24					
Female		2 / 3 (67%)	4 / 7 (57%)	4 / 7 (57%)	3 / 7 (43%)	13 / 24 (54%)
Male		1 / 3 (33%)	3 / 7 (43%)	3 / 7 (43%)	4 / 7 (57%)	11 / 24 (46%)
Ethnicity	24					
White		3 / 3 (100%)	4 / 7 (57%)	4 / 7 (57%)	5 / 7 (71%)	16 / 24 (67%)
Mixed/multiple groups		0 / 3 (0%)	1 / 7 (14%)	0 / 7 (0%)	0 / 7 (0%)	1 / 24 (4.2%)
Asian/Asian British		0 / 3 (0%)	2 / 7 (29%)	2 / 7 (29%)	1 / 7 (14%)	5 / 24 (21%)
Other ethnic group		0 / 3 (0%)	0 / 7 (0%)	1 / 7 (14%)	1 / 7 (14%)	2 / 24 (8.3%)
Fitzpatrick type	24					
I		0 / 3 (0%)	1 / 7 (14%)	1 / 7 (14%)	1 / 7 (14%)	3 / 24 (13%)
II		1 / 3 (33%)	1 / 7 (14%)	0 / 7 (0%)	1 / 7 (14%)	3 / 24 (13%)
III		1 / 3 (33%)	4 / 7 (57%)	2 / 7 (29%)	4 / 7 (57%)	11 / 24 (46%)
IV		1 / 3 (33%)	1 / 7 (14%)	4 / 7 (57%)	1 / 7 (14%)	7 / 24 (29%)
CMV (visit 1)	22					
Not detected		2 / 2 (100%)	4 / 7 (57%)	2 / 7 (29%)	4 / 6 (67%)	12 / 22 (55%)
Detected		0 / 2 (0%)	3 / 7 (43%)	5 / 7 (71%)	2 / 6 (33%)	10 / 22 (45%)
Unknown		1	0	0	1	2

[1] Median [Q1, Q3]; n / N (%)

Continuous variables are median [IQR]; categorical variables are n/N (%).

Table 4.3: KLH2: Demographics by randomisation group

Characteristic	N	Randomisation group		
		KLH 1000 + Alhydrogel N = 6	**KLH 1000 + Montanide** N = 6	**Overall** N = 12
Age (years)	12	23.0 [22.0, 25.0]	24.5 [19.0, 29.0]	23.5 [21.5, 27.0]
Height (cm)	12	171.0 [168.7, 176.0]	178.5 [174.0, 184.0]	175.0 [168.9, 179.3]
Weight (kg)	12	78.9 [57.0, 84.0]	79.4 [76.8, 97.0]	79.4 [68.9, 90.5]
BMI (kg/m^2)	12	25.2 [20.0, 29.4]	25.9 [22.7, 28.3]	25.9 [22.7, 28.8]
Sex	12			
Female		2 / 6 (33%)	2 / 6 (33%)	4 / 12 (33%)
Male		4 / 6 (67%)	4 / 6 (67%)	8 / 12 (67%)
Ethnicity	12			
Asian or Asian British		1 / 6 (17%)	3 / 6 (50%)	4 / 12 (33%)
Black, Black British, Caribbean or African		2 / 6 (33%)	1 / 6 (17%)	3 / 12 (25%)
Other ethnic group		3 / 6 (50%)	2 / 6 (33%)	5 / 12 (42%)
Fitzpatrick type	12			
I		0 / 6 (0%)	2 / 6 (33%)	2 / 12 (17%)
II		1 / 6 (17%)	0 / 6 (0%)	1 / 12 (8.3%)
III		1 / 6 (17%)	3 / 6 (50%)	4 / 12 (33%)
IV		3 / 6 (50%)	0 / 6 (0%)	3 / 12 (25%)
V		1 / 6 (17%)	0 / 6 (0%)	1 / 12 (8.3%)
VI		0 / 6 (0%)	1 / 6 (17%)	1 / 12 (8.3%)

[1] Median [Q1, Q3]; n / N (%)

Continuous variables are median [IQR]; categorical variables are n/N (%).

4.3.2 Humoral response to primary KLH immunisation, and the effect of adjuvants

Both adjuvanted regimens elicited robust humoral responses significantly exceeding those observed with unadjuvanted KLH. In the pairwise analysis of the log_{10} IgG AUC_{0-28}, both Alhydrogel (Mean Difference = 16.4 $\log_{10}$U·days, 95% CI 7.9 to 24.9, $p < 0.001$) and Montanide (Mean Difference = 14.3 $\log_{10}\text{U} \cdot \text{days}$, 95% CI 5.79 to 22.8, $p = 0.001$) were statistically superior to KLH alone. However, no significant difference was observed between the two adjuvants (Mean Difference = -2.11, $p = 0.521$). This finding was confirmed by the ANCOVA model adjusting for age, sex, and CMV status ($p = 0.57$). The percentage of participants seroconverting ($\geq$ 3 s.d. above mean baseline) at Day 28 was 0% for both the placebo and KLH only groups (0/3 and 0/7 participants respectively), compared to 86% (6/7 participants) in both adjuvanted KLH groups. Thus both adjuvanted formulations enhanced humoral priming, with no discernable differences between Montanide and Alhydrogel.

4.3.3 T cell (ELISpot) response to primary KLH immunisation

KLH-specific IFNγ and IL-4 responses showed parallel kinetics, with responses detectable as early as Day 7, peaking at Day 14, and remaining detectable through Day 28 (Figure 4.8, Tables 4.8 and 4.9). In contrast to the serological findings, the magnitude of the cellular response was highly sensitive to the choice of adjuvant. Montanide ISA-51 generated significantly greater total T cell exposure (AUC_{0-28}) compared to Alhydrogel.For IFNγ, the pairwise comparison revealed a significant increase in the Montanide group relative to Alhydrogel (Mean Difference = 27,300 SFC·days, 95% CI 3,560 to 51,000, $p = 0.015$). A similar pattern was observed for IL-4 responses, where Montanide again elicited significantly higher AUC values compared to Alhydrogel (Mean Difference = 8,430 SFC·days, 95% CI -483 to 17,300, $p = 0.044$). While Alhydrogel elicited responses numerically higher than unadjuvanted KLH ($p = 0.093$), this difference was less pronounced than that observed with Montanide ($p = 0.003$). These data demonstrate that Montanide is significantly superior to Alhydrogel for the induction of systemic T cell responses in this model.

4.3.4 Effect of biological covariates on systemic responses

We explored the potential influence of biological covariates on the immune response using ANCOVA. In this small cohort (n = 10–13 complete cases), neither sex nor baseline CMV serostatus was a statistically significant predictor of the magnitude of the response for any outcome measure ($p > 0.05$ for all comparisons). For example, in the analysis of IFNγ AUC, while CMV seropositivity was associated with a moderate effect size (Partial $\eta^2 = 0.36$), this did not reach statistical significance ($F_{1,5} = 2.82$, $p = 0.154$).

4.3.5 Effect of intradermal KLH rechallenge dose on peak (48 h) clinical response

Induration and erythema Intradermal KLH rechallenge elicited dose-dependent induration and erythema, with response rate and magnitude most pronounced among participants receiving primary immunisation with KLH + Montanide (Figure 4.9, table 4.16). For example, at the 10 µg rechallenge dose, response rates (induration diameter ≥5mm) were 0% (0/3), 57% (4/7), 71% (5/7), and 100% (7/7) in the placebo, KLH only, KLH + Alhydrogel, and KLH + Montanide groups respectively. Magnitude of responses were similarly stratified, with geometric mean responses of 1.5 mm (95% CI 0.0-4.4mm), 7.9mm (95% CI 4.5 - 11.2 mm), 13.8 mm (95 % CI 4.0 - 23.5), and 23.1 mm (95% CI 13.8 - 32.4 mm) in the placebo, KLH only, KLH + Alhydrogel, and KLH + Montanide groups at the 10 µg rechallenge dose respectively. Erythema response showed a similar pattern (Figure 4.9, table 4.16).

Emax dose-response modelling We applied a Bayesian E_{max} framework to formally model the dose-induration data. Visual inspection of the individual dose-response curves suggested heterogeneity, with some participants showing minimal or absent responses across all doses. To account for this, data were filtered to exclude "non-responders" (defined as maximum induration <5 mm across all doses), focusing the model on the dose-response relationship within the sensitised population.

Model diagnostics and posterior predictive checks indicated a satisfactory fit to the data. The

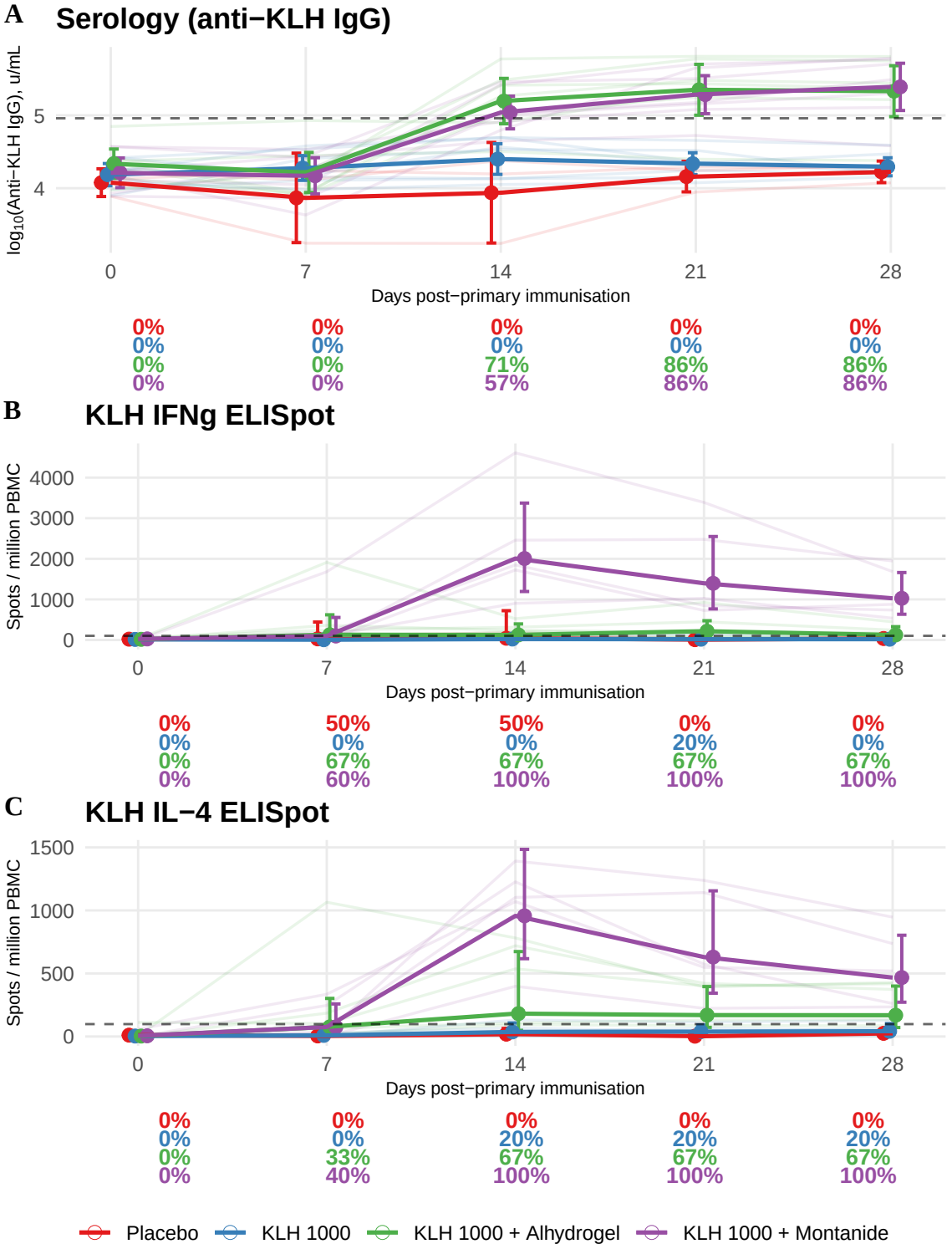

Figure 4.8: Time course of immune responses in peripheral blood following primary KLH immunisation, stratified by randomised study group. (A) Serological responses measured by anti-KLH ELISA (units/mL); values are $\log_{10}$-transformed. (B) KLH-specific IFNγ ELISpot responses (SFC per 10^6 PBMCs). (C) KLH-specific IL-4 ELISpot responses (SFC per 10^6 PBMCs). The horizontal dashed line indicates the positivity threshold defined as three standard deviations above the mean baseline (Day 0) response. Numbers below each panel indicate the percentage of participants classified as responders at each timepoint within each group.

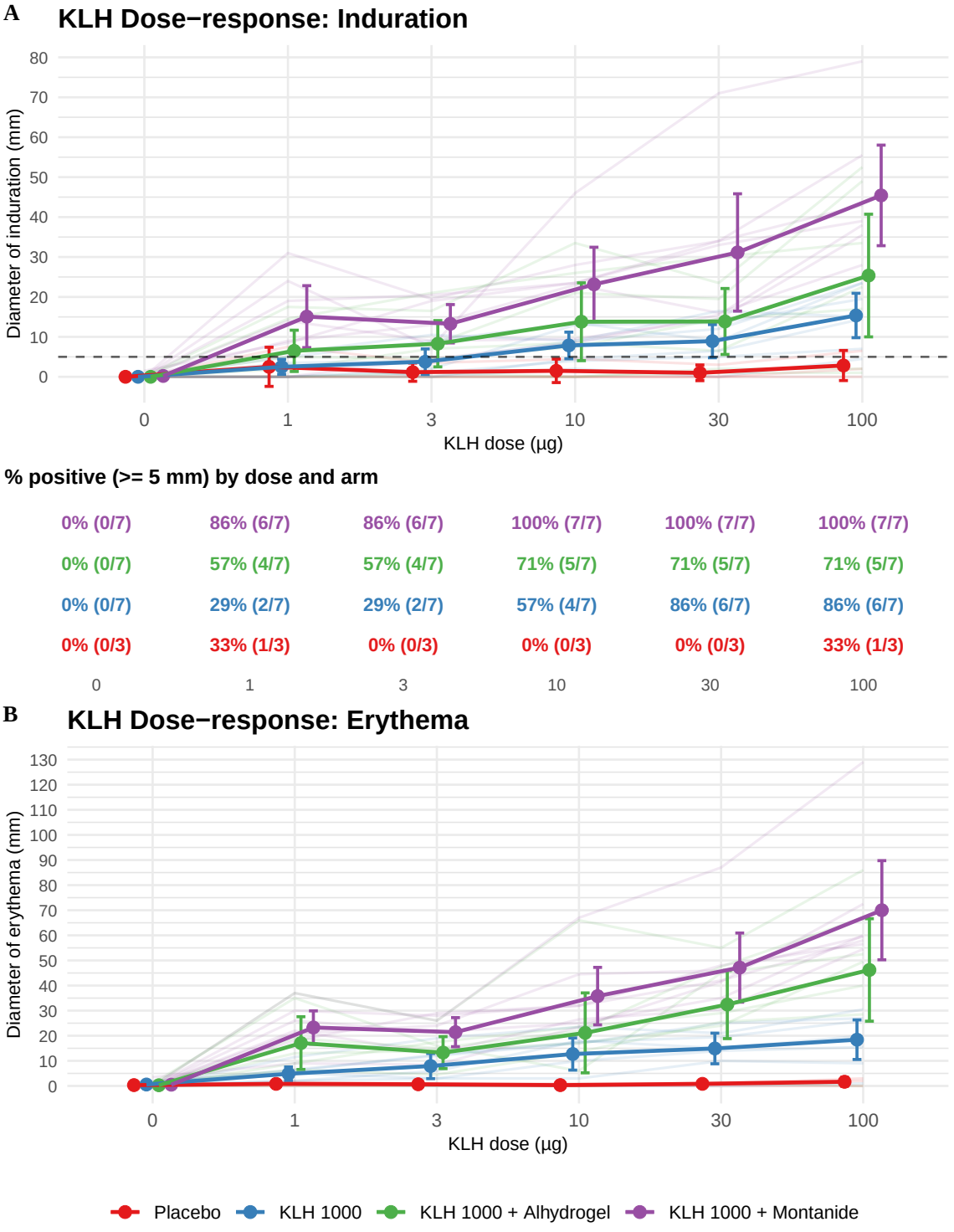

Figure 4.9: KLH rechallenge dose–response relationship for clinical cutaneous endpoints. (A) Mean diameter of induration (mm) as a function of intradermal KLH rechallenge dose (μg). (B) Mean diameter of erythema (mm) as a function of intradermal KLH rechallenge dose (μg). Light-coloured lines show individual participant trajectories, while bold lines indicate group-level mean responses. Points and error bars represent group means and 95% confidence intervals at each dose level.

resulting posterior distributions confirmed a sigmoidal dose-response relationship approaching a plateau at approximately 30–100 μg KLH (Figure 4.11). Posterior estimates for the maximum theoretical (mean) response (E_{max}) differed substantially by primary immunisation group. Montanide was associated with the highest maximal response, with a high probability of exceeding the maximal response of the Alhydrogel group ($Pr(E_{max}^{Mont} > E_{max}^{Alum}) = 0.89$). In contrast, the potency of the response (represented by the ED_{50}) was relatively conserved across adjuvanted groups. There was no strong evidence that Montanide lowers the sensitisation threshold compared to Alhydrogel ($Pr(ED_{50}^{Mont} < ED_{50}^{Alum}) = 0.45$). This suggests that while adjuvants enhance the magnitude of the effector response (E_{max}), the sensitivity to antigen (ED_{50}) is likely an intrinsic property of the KLH antigen itself once sensitisation has occurred.

Defining an optimal rechallenge dose Using the fitted model, we performed posterior predictive simulations to identify a standard rechallenge dose for future protocols. We defined an "optimal" response window as an induration diameter between 10–60 mm; this lower bound ensures feasibility for 4–6 mm punch biopsies, while the upper bound avoids saturation (insensitivity to modulation) and minimises risk of dose-related local adverse events. Simulations indicated that a 10 μg rechallenge dose was optimal. At this dose, the probability of achieving the target response window (10–60 mm) was 90% for KLH + Alhydrogel and 77% for KLH + Montanide. In contrast, a 1 μg dose yielded lower probabilities (53% and 76% respectively), primarily due to a higher risk of sub-10 mm responses. Based on these operating characteristics, 10 μg was selected as the standard rechallenge dose for the subsequent time-course cohort (KLH2).

4.3.6 Compositional analysis of skin biopsy cell populations

To characterise the cellular characteristics of the KLH response, we analysed flow cytometry data from 48 h punch biopsies using Bayesian compositional models. We first established the cellular correlates of the clinical response.

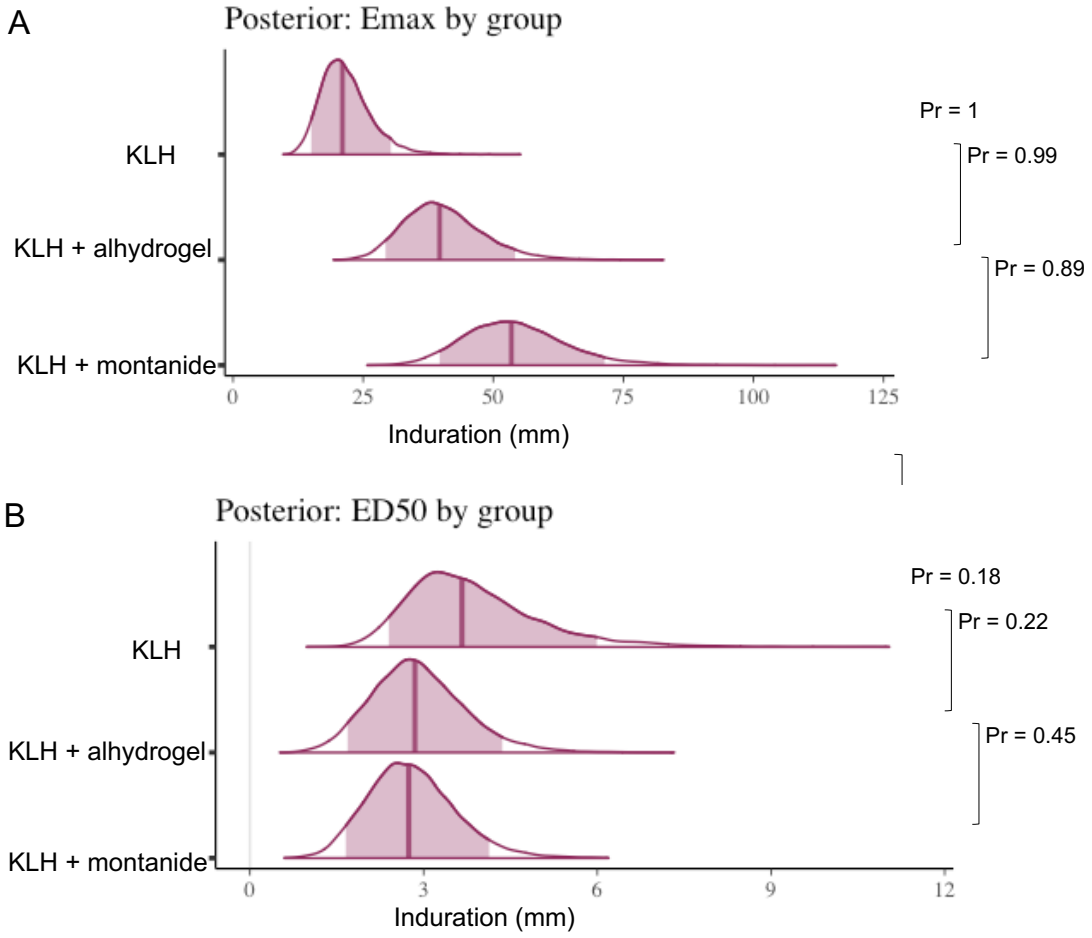

Figure 4.10: Posterior distributions of population-level parameters derived from the Bayesian E_{max} dose–response model. (A) Posterior density estimates for the maximal induration response (E_{max}, mm), stratified by primary adjuvant regimen. Primary sensitisation with Montanide is associated with a higher maximal response compared with Alhydrogel and unadjuvanted KLH. (B) Posterior density estimates for the half-maximal effective dose (ED_{50}, μg).

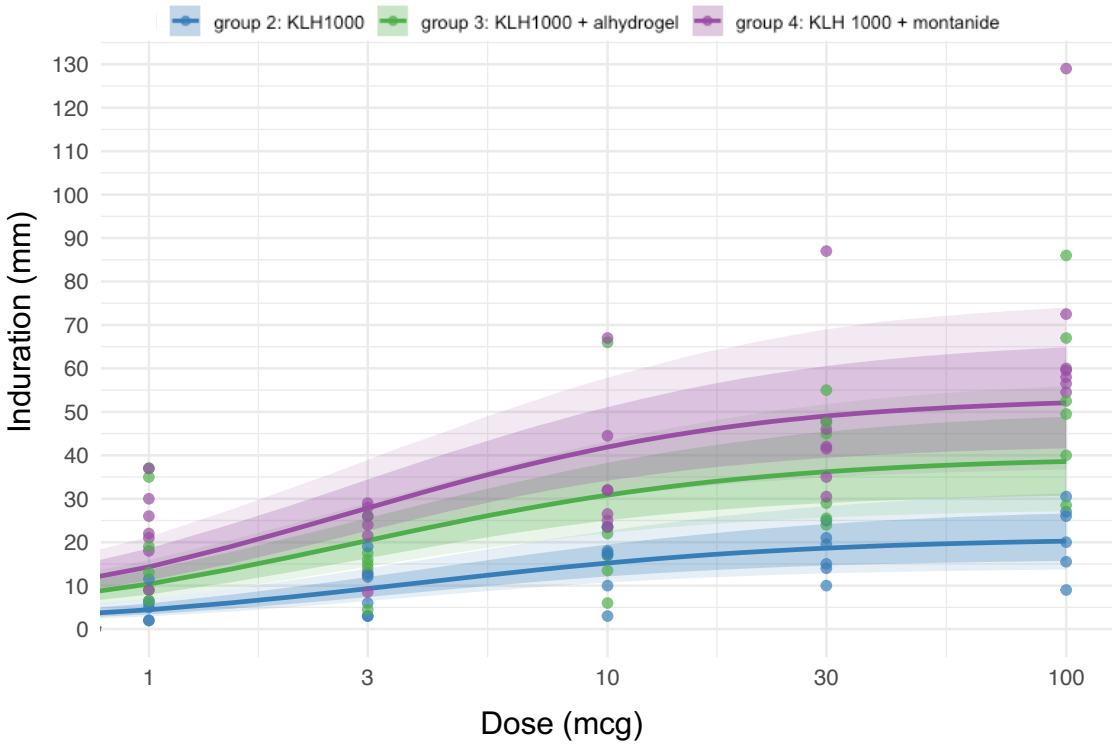

Figure 4.11: Modelled relationship between intradermal KLH rechallenge dose and diameter of clinical induration (mm) at 48 h. Solid lines show the mean posterior predictive dose–response for each primary immunisation group, while shaded ribbons denote the 50% (inner) and 90% (outer) credible intervals.

Cellular correlates of induration Comparison of indurated (≥5 mm) versus non-indurated sites confirmed activation of multiple immune axes consistent with a rechallenge response. Indurated lesions were characterised by influx of myeloid and lymphoid cells, with significant increases in the proportions of monocytes (classical, intermediate, and non-classical phenotypes), $CD4^+$ T cells and B cells relative to resident populations (Figure 4.13). When induration was modelled as a continuous variable (Model 2), the magnitude of the clinical response was positively associated with the relative abundance of $CD4^+$ T cells, monocytes, macrophages and conventional dendritic cells, suggesting that clinical palpation is a valid proxy for the magnitude of cellular infiltration.

Effect of rechallenge dose and adjuvant on cellular composition We next assessed the impact of rechallenge dose (1 μg vs 100 μg) and primary adjuvant (Montanide vs Alhydrogel) on the composition of the infiltrate (Model 3). Consistent with the clinical dose-response modelling, increasing the antigen dose from 1 μg to 100 μg resulted in a significant shift in composition, driven primarily by an enrichment of myeloid populations. Notably, there was no associated increase in lymphoid populations identified, suggesting that the higher dose primarily drove a non-specific (myeloid) inflammatory response rather than a classical adaptive immune response. (Figure 4.15). Targeted analysis of the $CD4^+$ T cell compartment (Model 4a, 4b) revealed a mixed phenotype at 48 h, with the higher rechallenge dose eliciting a greater proportion of proliferating ($Ki67^{+)}$ $CD4^+$ T cells and reduced proportions of Treg cells. Higher rechallenge dose was also associated with alterations of transcription factor expression, most notably a higher proportion of $CD4^+TBET^+CXCR3^+$ T cells (Th1) (Figure 4.17).

In contrast to the clear dose-effects observed, the choice of primary adjuvant had a more limited impact on the cellular composition of the recall response at 48 h. After controlling for dose, with small but statistically significant increases in $CD4^+$ T cells and conventional dendritic cells, and corresponding decreases in macrophages, NK cells, and NK-T cells in participants receiving Alhydrogel-adjuvanted primary sensitisation versus those receiving Montanide. (Figure 4.15). Classical $CD4^+$ T cell phenotypes were similar across adjuvant groups with greater proportions of naive T cells and reduced proportions of $CD4^+$ T cells expressing PD1 in the Alhydrogel group. Montanide

was associated with a greater influx of $CD4^+TBET^+CXCR3^+$ T cells, however there was not a corresponding enrichment of $CD4^+GATA3^+$ T cells therefore suggesting a more complex effect of primary adjuvant on Th1/Th2 polarisation than hypothesised *a priori*.

4.3.7 Temporal evolution of the cutaneous immune response

While the dose-response analysis focused on the classical 48 h DTH window, the time-course cohort (KLH2) revealed distinct kinetic profiles for key immune populations, demonstrating that the qualitative nature of the immune response is highly dependent on the sampling timepoint. In contrast, inclusion of primary adjuvant group as a covariate in the time-course model was not statistically significant suggesting temporal profiles following rechallenge are not substantially altered by the adjuvants assessed.

Absolute cell counts While a compositional approach represents a robust method of accounting for the practical challenges of absolute quantification of cell numbers in biopsy specimens, observation of absolute numbers provides important context as to the magnitude of the immune response observed across conditions. Figure 4.12 demonstrates clearly the substantial immune cell influx (absolute counts normalised to biopsy area)associated with active challenge sites compared to controls, with an overall peak infiltrate seen at 48 h, declining marginally at five days and more substantially by 14 days. Visual inspection of this figure identifies independent temporal trajectories: For monocytes, an abrupt decline is seen between 48 h to 5 days, with more modest declines for $CD4^+$ T cells and increases for $CD8^+$ T cells between these timepoints. For B cells, a modest increase is seen between 48 h and 5 days, declining at 14 days. The consequences of the relative trajectories of each of these cell types is evaluated formally in the compositional analysis.

Comparison of control conditions The temporal compositional analysis included two distinct control groups. For the KLH1 study, control biopsies were taken from a site receiving 0.9% saline injection 48 h prior. In contrast, for the KLH2 study, control biopsies were taken from unchallenged skin. Controls appeared to differ systematically in composition; KLH2 control biopsies

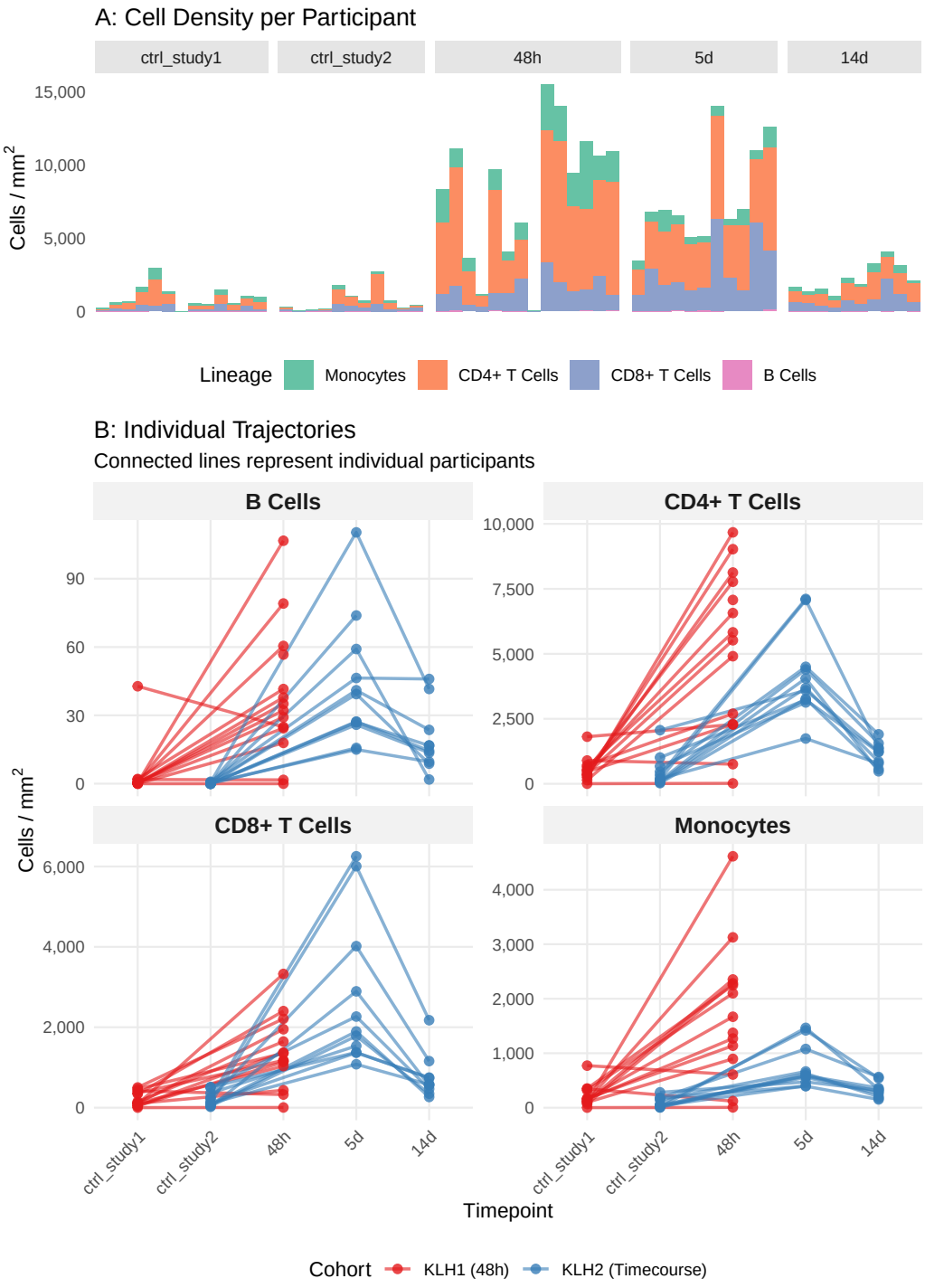

Figure 4.12: Time course of major immune cell populations based on normalised absolute cell counts. Absolute counts were normalised to biopsy surface area to account for differences in tissue processing between studies. In KLH1, flow cytometry was performed on one half of a 5 mm punch biopsy (approximate area 39.2 mm^2), whereas in KLH2, one quarter of a 6 mm punch biopsy was analysed (approximate area 28.2 mm^2). For KLH2, cell counts represent the geometric mean across the 1 μg and 100 μg rechallenge doses. (A) Cell density over time by population (monocytes, $CD4^+$ T cells, $CD8^+$ T cells, and B cells). (B) Individual participant trajectories showing timepoint versus cell density, coloured by study (KLH1: red; KLH2: blue).

demonstrated modest increases in most immune cell populations relative to KLH1 controls, and a corresponding relative decrease in CD45$^-$ (non-immune) cells. The causes of this are unclear, although it is plausible that the biopsy compositions were affected by local or systemic rechallenge dynamics. Thus, their status as true 'controls' requires further interrogation.

Temporal shift in dominant myeloid and lymphoid populations At 48 h, the cellular infiltrate was heterogeneous and inflammatory, characterised by a dominant influx of classical monocytes ($CD14^+CD16^-$) and naive CD4$^+$ T cells, consistent with rapid recruitment from the circulation. By Day 5 and Day 14, the infiltrate became increasingly lymphocyte-predominant with progressive accumulation of both CD4$^+$ and CD8$^+$ T cells (Figure 4.20 and 4.21). In addition, we observed a dynamic shift within the monocyte compartment, with progressive enrichment of intermediate ($CD14^+CD16^+$) and non-classical monocytes ($CD14^-CD16^+$) from 48 h through to Day 14. While B cells were a minor component at all timepoints, with absolute counts peaking around 48 h to 5 days (Figure 4.12), the persistence relative to other immune cell types (hence proportional enrichment seen in the compositional analysis, Figure 4.20) suggests independent kinetic determinants, a finding that warrants histological investigation in future studies.

Temporal skewing of T helper phenotypes The phenotypic profile of the infiltrating CD4$^+$ T cells also exhibited significant temporal evolution (Figure 4.21. At 48 h the CD4$^+$ populations were characterised by a high proportion of naive cells, while we noted a striking peak in proliferating (Ki67$^+$) CD4$^+$ T cells at day 5 which had largely returned to baseline by day 14. We noted a progressive increase in CD4$^+$ resident memory T cells between 48 h and day 14 although these remained below control proportions due to a corresponding rise in PD1$^+$ CD4$^+$ T cells. Perhaps unexpectedly, we did not notice an accumulation of regulatory T cells over time.

Early timepoints (48 h) displayed a mixed T_h1/T_h2 profile with balanced expression of the transcription factors T-bet and GATA3. However, by Day 14, cells expressing T-bet were significantly more prominent, indicating a progressive polarization toward a T_h1 phenotype in the resolving phase of the response (Figure 4.22). This suggests that the local inflammatory milieu drives a Th1 shift

over time, regardless of the primary adjuvant used.

4.3.8 Transcriptomic profile of cutaneous response 48 h following intradermal rechallenge

Due to sample yield limitations in the control and low-dose sites, bulk RNA-seq analysis was restricted to 48 h biopsy samples from active rechallenge sites (100 μg) in sensitised individuals. Notably, despite the clear differences in circulating T cell magnitude observed by ELISpot, differential expression analysis and PCA visualisation demonstrated no differences according to primary adjuvant (Figure 4.18 B and F). We utilised Singscore to assess pathway-level activity on a single-sample basis.As expected, KLH rechallenge elicited a strong transcriptional signature consistent with an active DTH response. We observed significant enrichment of gene sets associated with antigen processing and presentation, TNFα signaling via NFκB, downstream TCR signaling, and IL-6/JAK/STAT signaling (Figure 4.19). Interferon alpha/beta signaling pathways were also highly enriched. Together these results suggest that while the magnitude of the peripheral response differs, the qualitative nature of the transcriptional program engaged at the tissue site at 48 h does not differ markedly across adjuvants, within the notable limitation of low statistical power in the setting of small sample sizes.

4.3.9 Agreement between clinical assessment and non-invasive imaging

To determine if objective imaging modalities could serve as surrogates for clinical assessment, we compared manual measurement of induration diameter with metrics derived from Laser Doppler Imaging (LDI) and Antera 3D multispectral imaging. A comparison of imaging modalities for a representative participant is shown in Figure 4.23.

Concordance between modalities Imaging metrics demonstrated strong positive correlations with clinical endpoints across all dose levels. Metrics quantifying extent of inflammation (clinical erythema, LDI increased flux mean diameter) were highly correlated with induration. Met-

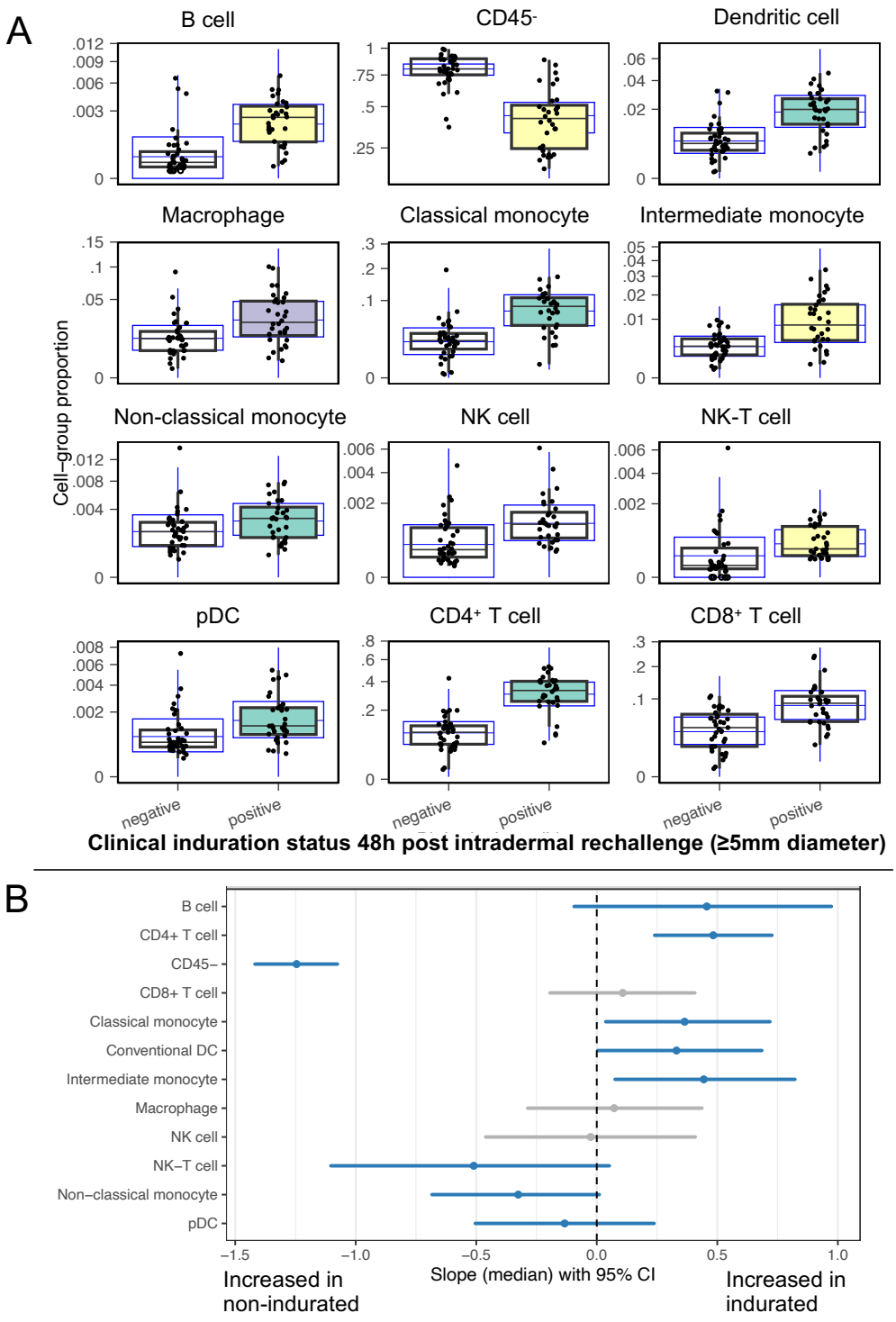

Figure 4.13: Model 1: Association between clinical induration status and the proportion of major immune cell populations 48 h after intradermal KLH rechallenge, assessed by flow cytometry. Clinical induration status was defined as negative (diameter <5 mm) or positive (diameter ≥5 mm). (A) Observed data (black boxplots) overlaid with posterior predictive distributions (blue boxplots), including individual observations (points) and fitted model estimates. Shaded fill indicates a statistically significant association (5% FDR) between induration status and the expected composition (green), variability in composition (purple), or both (yellow). (B) Posterior 95% credible intervals for the association between induration status and cell-type proportion on the logit scale, coloured by significance at 5% FDR (blue = significant; grey = non-significant).

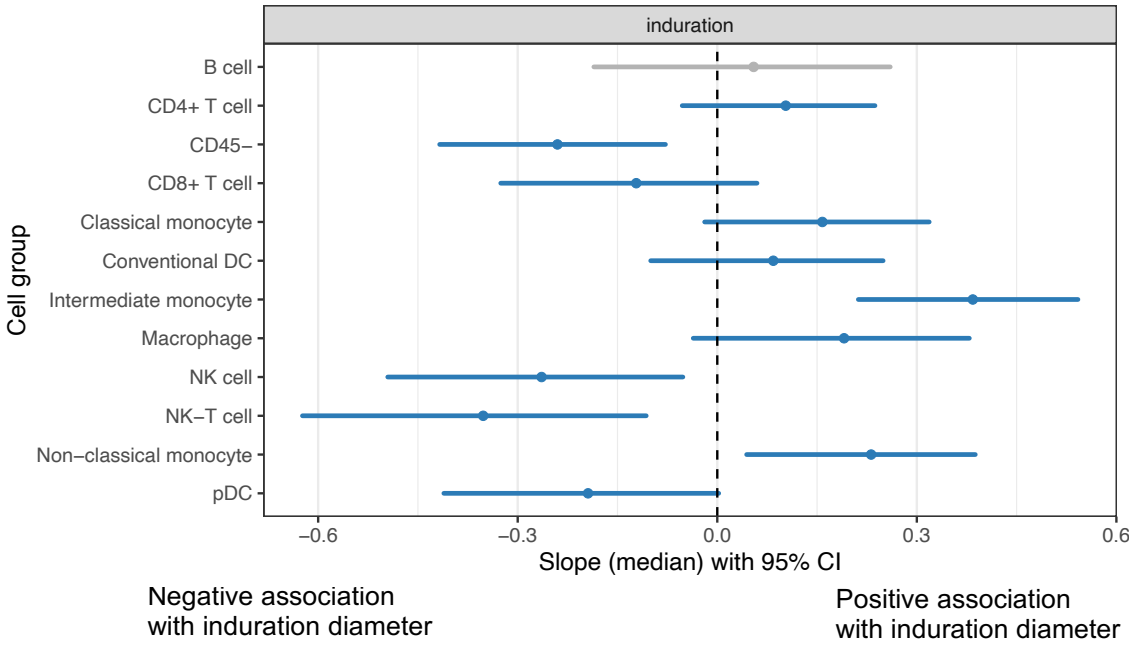

Figure 4.14: Model 2: Association between induration diameter and the proportion of major immune cell populations 48 h after intradermal KLH rechallenge, assessed by flow cytometry. Induration diameter was modelled as a continuous predictor. Points represent posterior mean estimates, with horizontal bars indicating posterior 95% credible intervals on the logit scale. Intervals are coloured by significance at a 5% FDR (blue = significant; grey = non-significant). Analysis includes only clinically indurated lesions (diameter $\geq$5 mm).

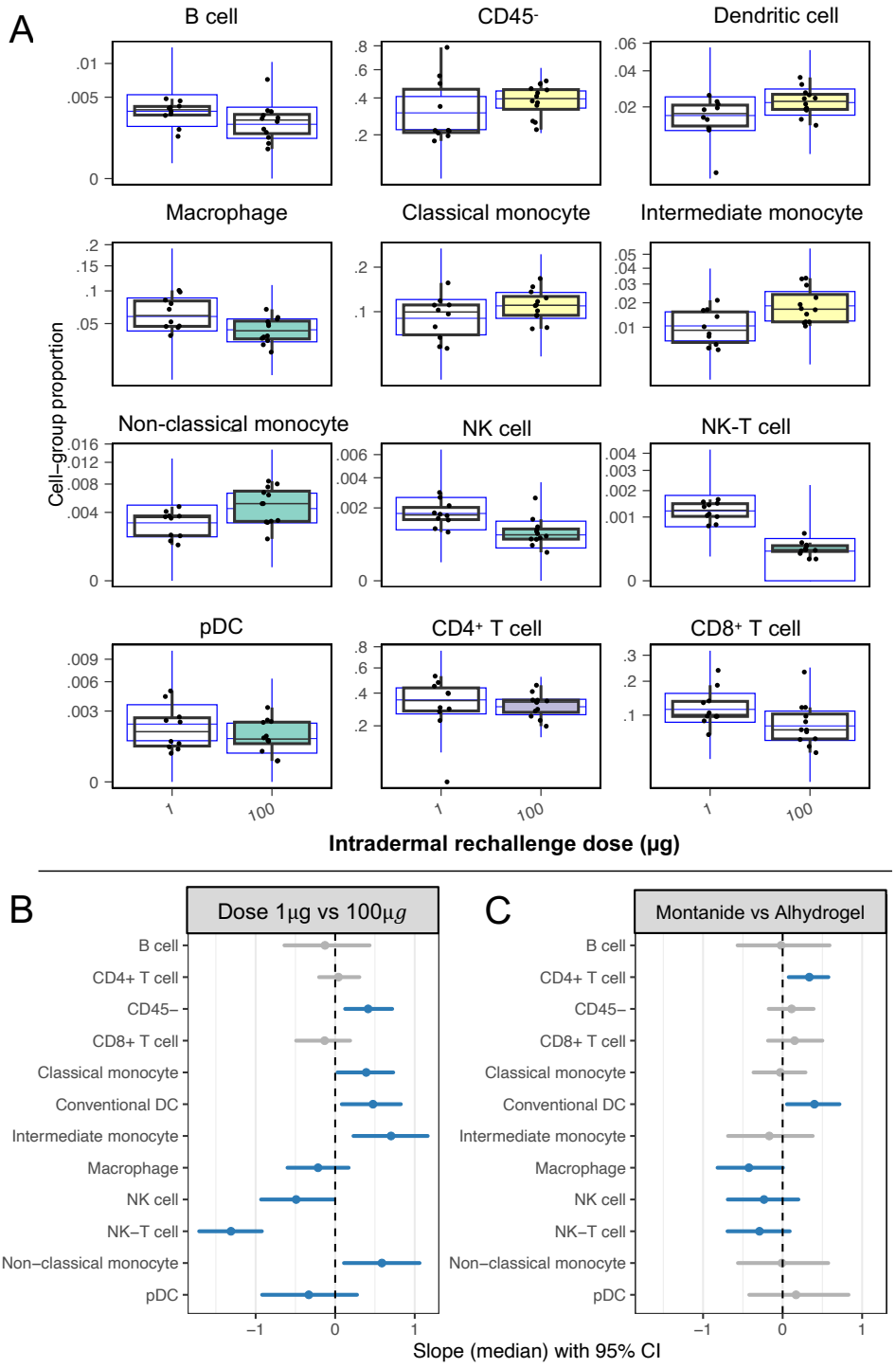

Figure 4.15: Model 3: Association between intradermal KLH rechallenge dose (1 μg vs 100 μg), primary adjuvant regimen (KLH + Alhydrogel versus KLH + Montanide), and the proportion of major immune cell populations 48 h after rechallenge, assessed by flow cytometry. (A) Observed data (black boxplots) overlaid with posterior predictive distributions (blue boxplots), including individual observations (points) and fitted model estimates. Shaded fill indicates a statistically significant association (5% FDR) between dose and the expected composition (green), variability in composition (purple), or both (yellow), controlling for primary adjuvant regimen. (B) Posterior 95% credible intervals for the association between rechallenge dose and cell-type proportion on the logit scale. (C) Posterior 95% credible intervals for the association between primary adjuvant regimen and cell-type proportion on the logit scale. Credible intervals are coloured by significance at 5% FDR (blue = significant; grey = non-significant).

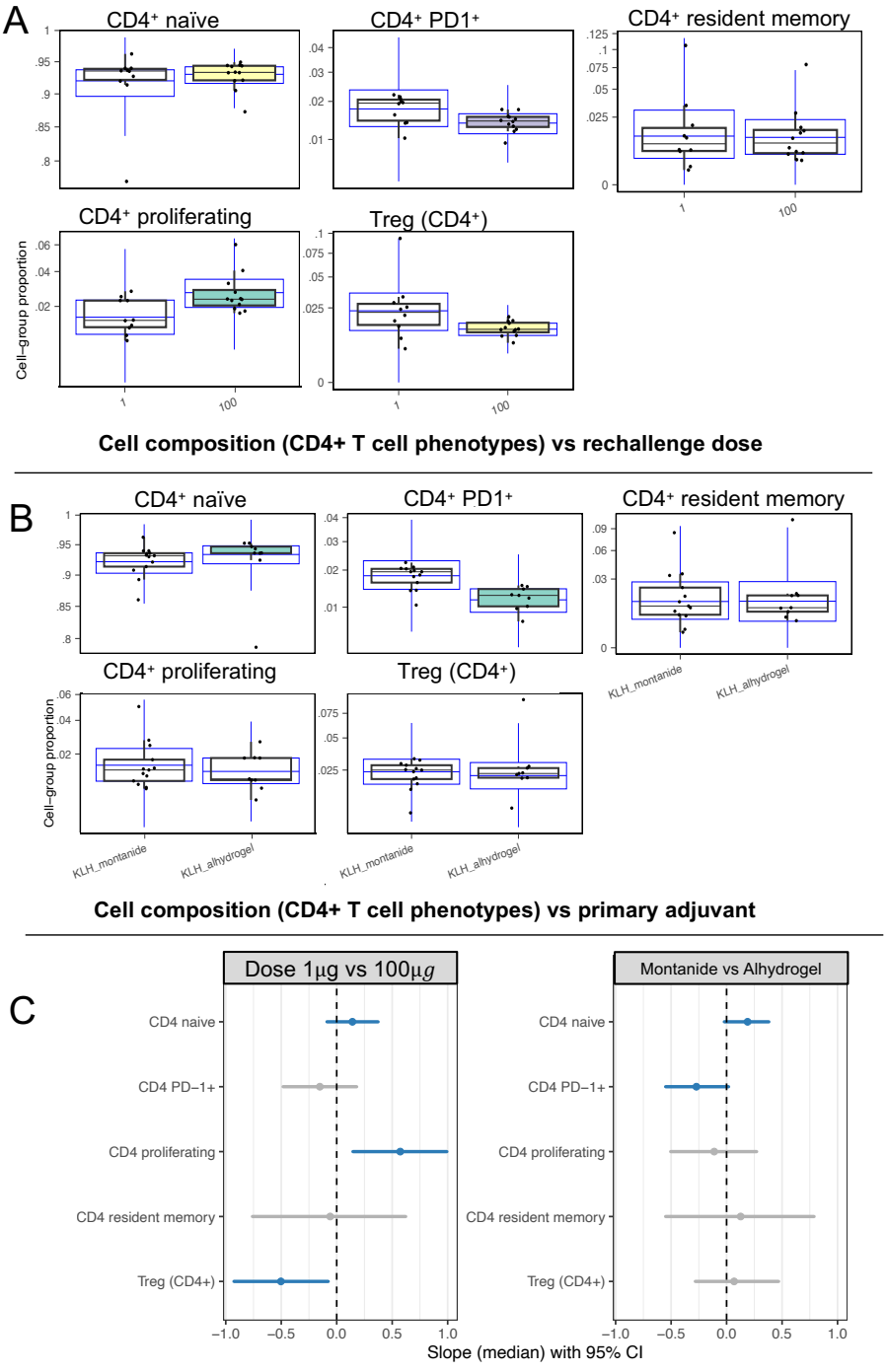

Figure 4.16: Model 4a: Association between intradermal KLH rechallenge dose (1 μg vs 100 μg), primary adjuvant regimen (KLH + Alhydrogel versus KLH + Montanide), and the proportion of CD4$^+$ T cell phenotypes 48 h after rechallenge, assessed by flow cytometry. (A) Observed data (black boxplots) overlaid with posterior predictive distributions (blue boxplots), including individual observations (points) and fitted model estimates. Shaded fill indicates a statistically significant association (5% FDR) between rechallenge dose and the expected composition (green), variability in composition (purple), or both (yellow), controlling for primary adjuvant regimen. (B) Posterior 95% credible intervals for the association between rechallenge dose and CD4$^+$ T cell phenotype proportion on the logit scale. (C) Posterior 95% credible intervals for the association between primary adjuvant regimen and CD4$^+$ T cell phenotype proportion on the logit scale. Credible intervals are coloured by significance at 5% FDR (blue = significant; grey = non-significant).

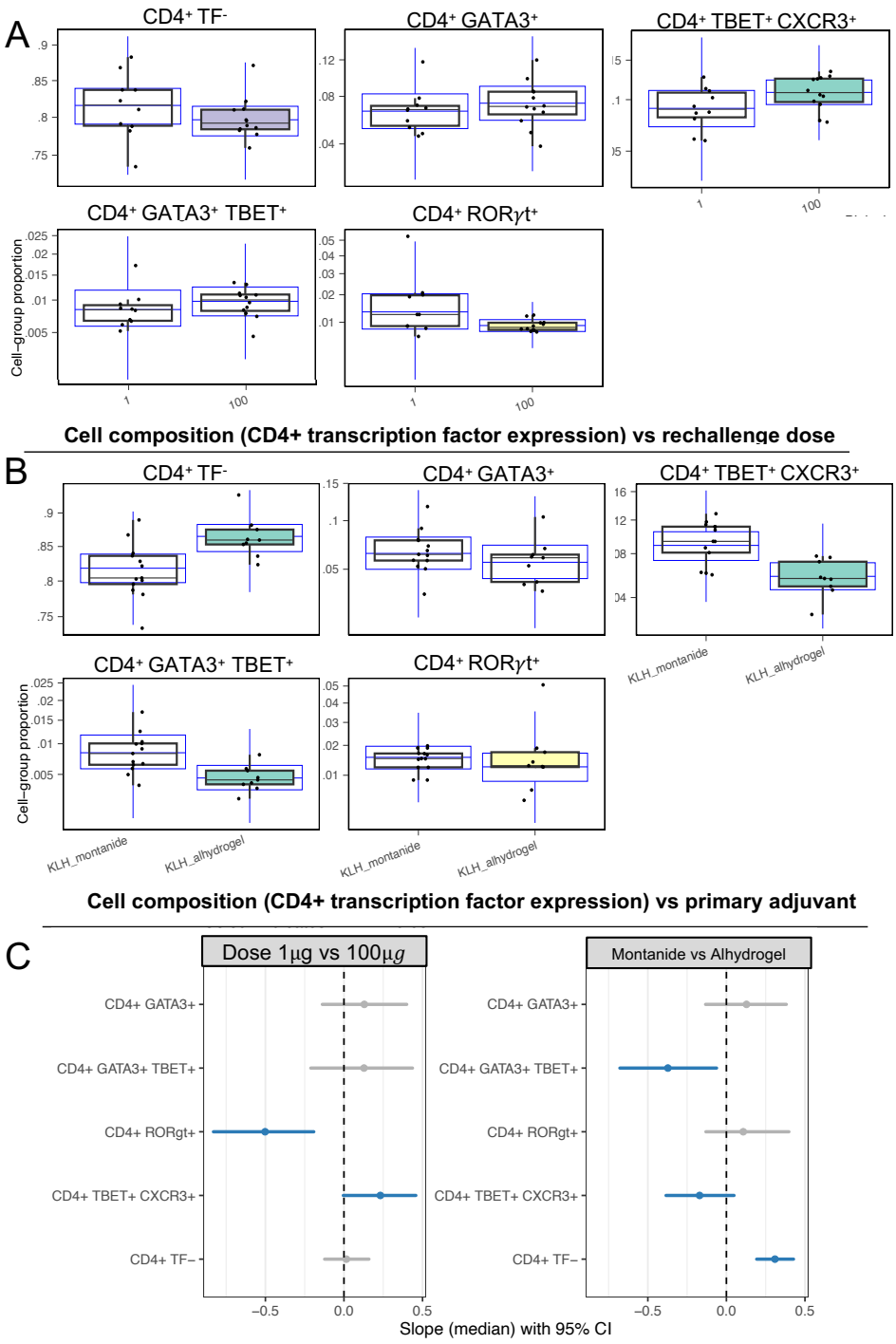

Figure 4.17: Model 4b: Association between intradermal KLH rechallenge dose (1 μg vs 100 μg), primary adjuvant regimen (Alhydrogel versus Montanide), and the proportion of CD4$^+$ T cells expressing lineage-defining transcription factors 48 h after rechallenge, assessed by flow cytometry. (A) Observed data (black boxplots) overlaid with posterior predictive distributions (blue boxplots), including individual observations (points) and fitted model estimates. Shaded fill indicates a significant association (5% FDR) between rechallenge dose and the expected composition (green), variability in composition (purple), or both (yellow), controlling for primary adjuvant regimen. (B) Posterior 95% credible intervals for the association between rechallenge dose and transcription factor–defined CD4$^+$ T cell subsets on the logit scale. (C) Posterior 95% credible intervals for the association between primary adjuvant regimen and transcription factor–defined CD4$^+$ T cell subsets on the logit scale. Credible intervals are coloured by significance at 5% FDR (blue = significant; grey = non-significant).

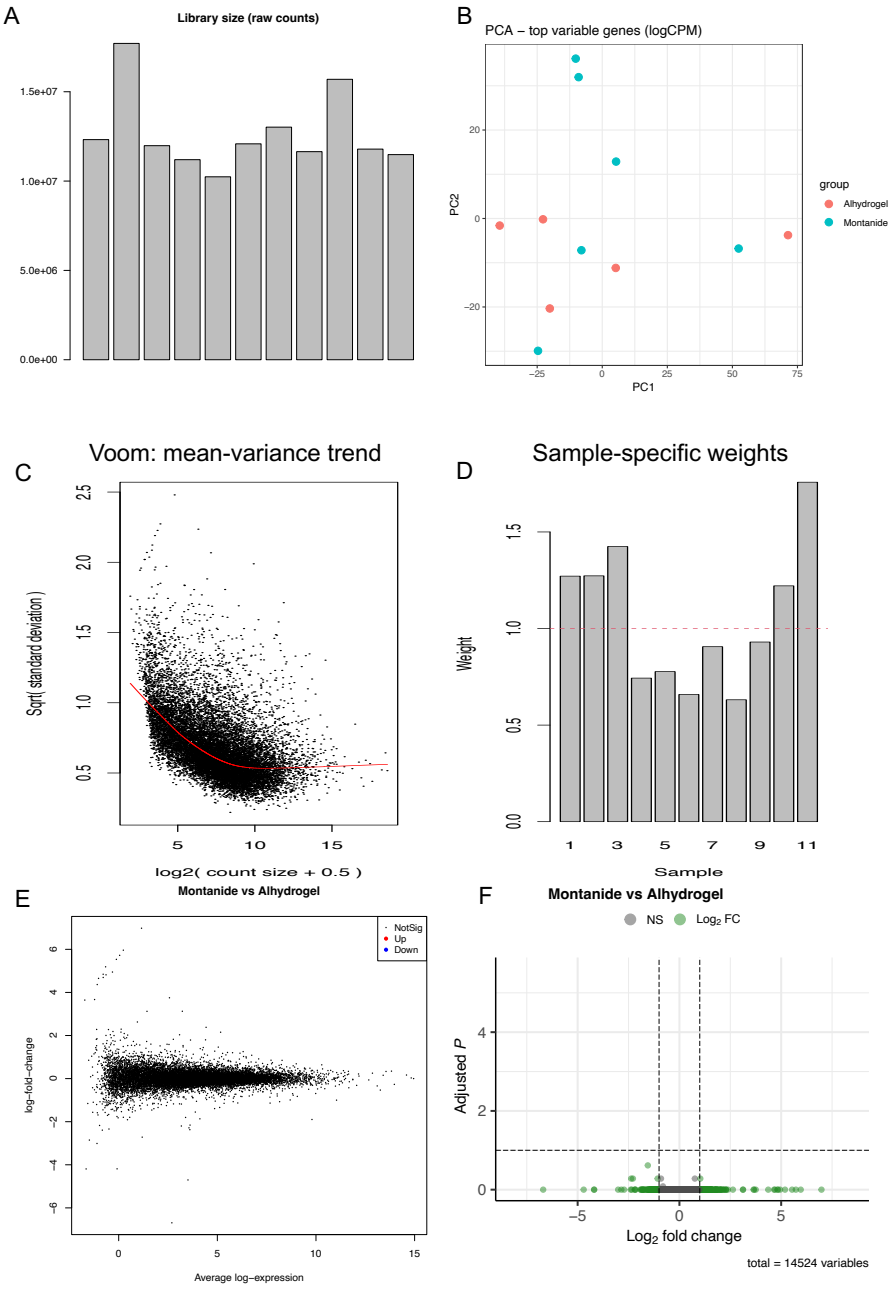

Figure 4.18: Bulk transcriptomic analysis of skin biopsies collected 48 h after intradermal KLH rechallenge. (A) Library sizes for RNA-seq samples included in the analysis ($n = 11$). (B) Principal component analysis (PCA) of voom-transformed gene expression values, coloured by primary adjuvant regimen (Alhydrogel vs Montanide). (C) Mean–variance trend estimated by the voom transformation. (D) Sample-specific precision weights estimated by voom. (E) Mean $\log_2$ expression versus $\log_2$ fold change for the linear model comparing primary adjuvant regimens. (F) Volcano plot showing $\log_2$ fold change versus Benjamini–Hochberg–adjusted p-value for the contrast between Alhydrogel- and Montanide-adjuvanted KLH. No genes met the significance threshold for differential expression between adjuvant groups.

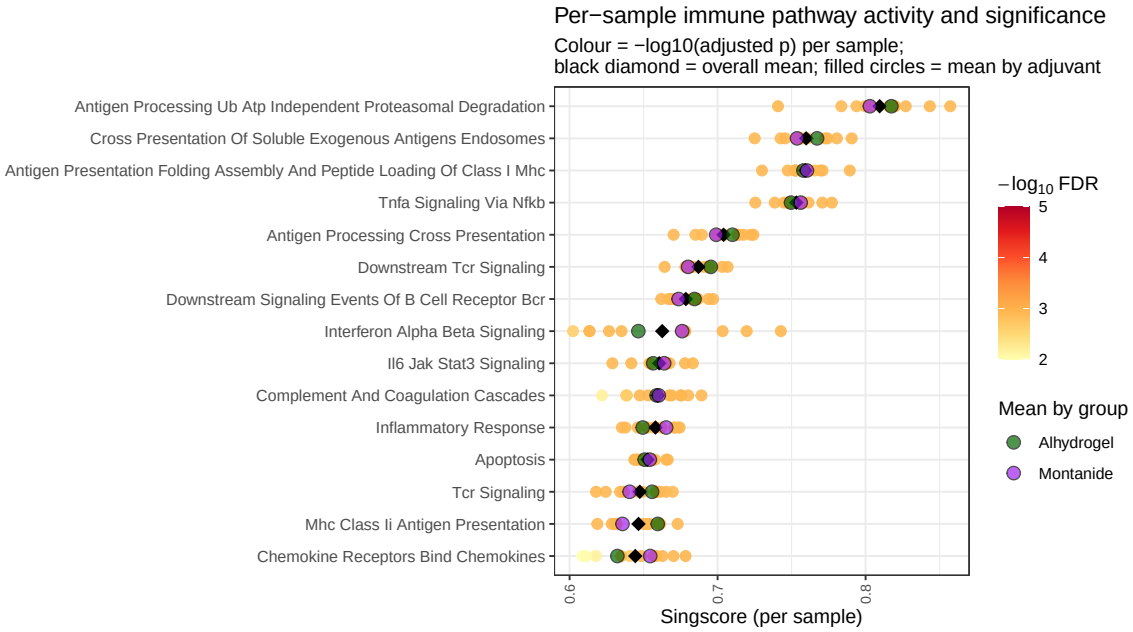

Figure 4.19: Exploratory single-sample gene set enrichment analysis of 48 h bulk RNA-seq data derived from enzymatically digested skin biopsy specimens following intradermal KLH rechallenge. RNA sequencing was performed only on highly cellular biopsy digests yielding $\geq$ 400,000 viable cells, from which a fixed aliquot of 100,000 cells was used for library preparation. Immune-related gene sets were selected from MSigDB, including Hallmark Immune Signatures (H collection), Reactome immune pathways (C2:CP:REACTOME), KEGG immune pathways (C2:CP:KEGG), and transcription factor target sets (C3:TFT_LEGACY). Each point represents the singscore for an individual sample; horizontal bars indicate the overall mean score across samples, with group-wise means shown for each primary adjuvant regimen. Points are coloured according to $-\log_{10}$ empirical p-value derived from permutation-based null distributions. Only the top 15 gene sets ranked by mean enrichment score are shown. This analysis was exploratory in nature and intended to inform the feasibility and optimisation of tissue processing and multi-modal readouts in future KLH challenge studies.

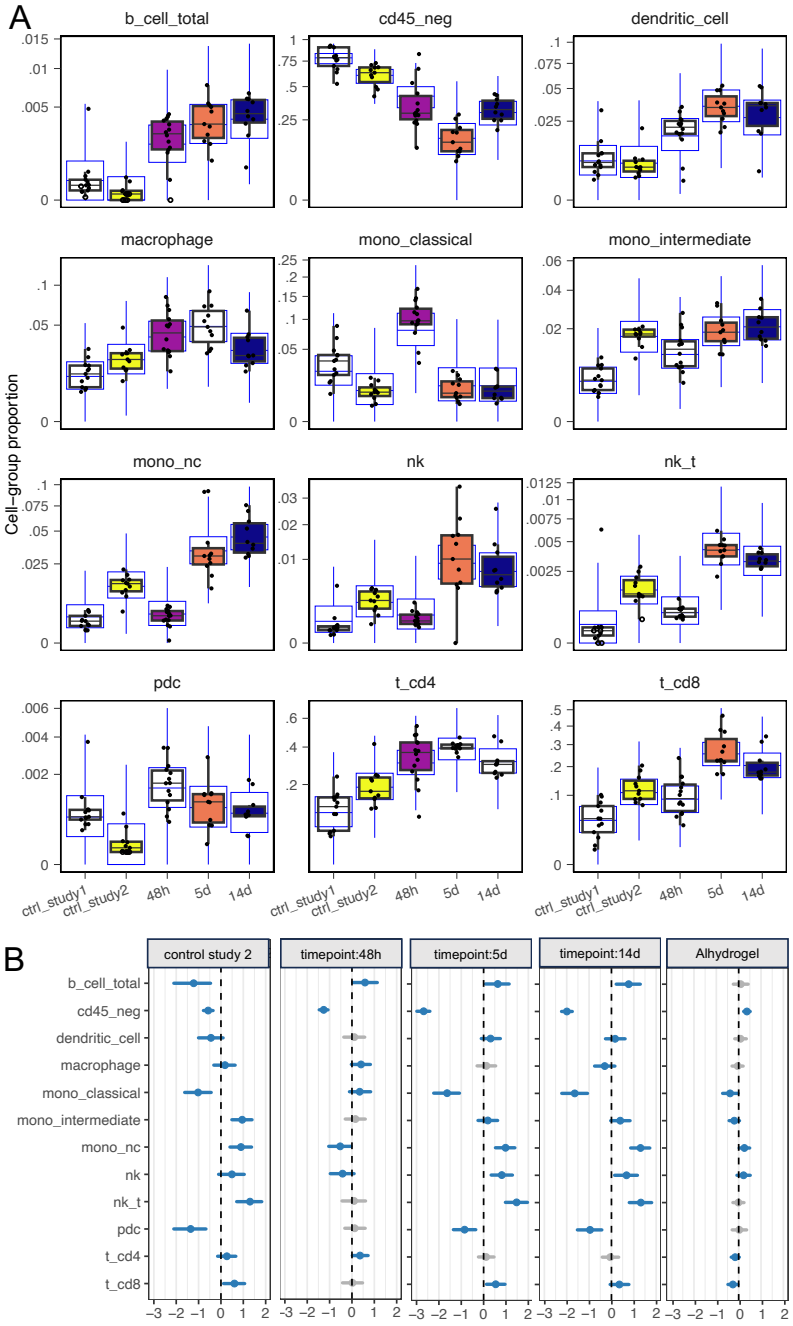

Figure 4.20: Compositional modelling results showing the temporal evolution of major immune cell populations following intradermal KLH rechallenge. Results are aggregated across the KLH1 and KLH2 studies. Cell-type proportions are modelled within each condition and timepoint using Bayesian compositional models. Statistical significance is assessed by comparison with the KLH1 control condition (ctrl_study1), which serves as the reference category for hypothesis testing.
(A) Boxplots summarising observed cell-type proportions at each timepoint. Boxplots are filled when the corresponding timepoint differs significantly from the KLH1 control at a 5% FDR, with distinct fill colours indicating different timepoints. (B) Interval plots showing posterior mean estimates and 95% credible intervals for the difference in expected cell-type proportion relative to the KLH1 control (logit scale) at each timepoint. The rightmost panel shows the effect of primary Alhydrogel adjuvantation relative to Montanide (reference). Intervals are coloured blue when the difference is statistically significant at a 5% FDR and grey otherwise.

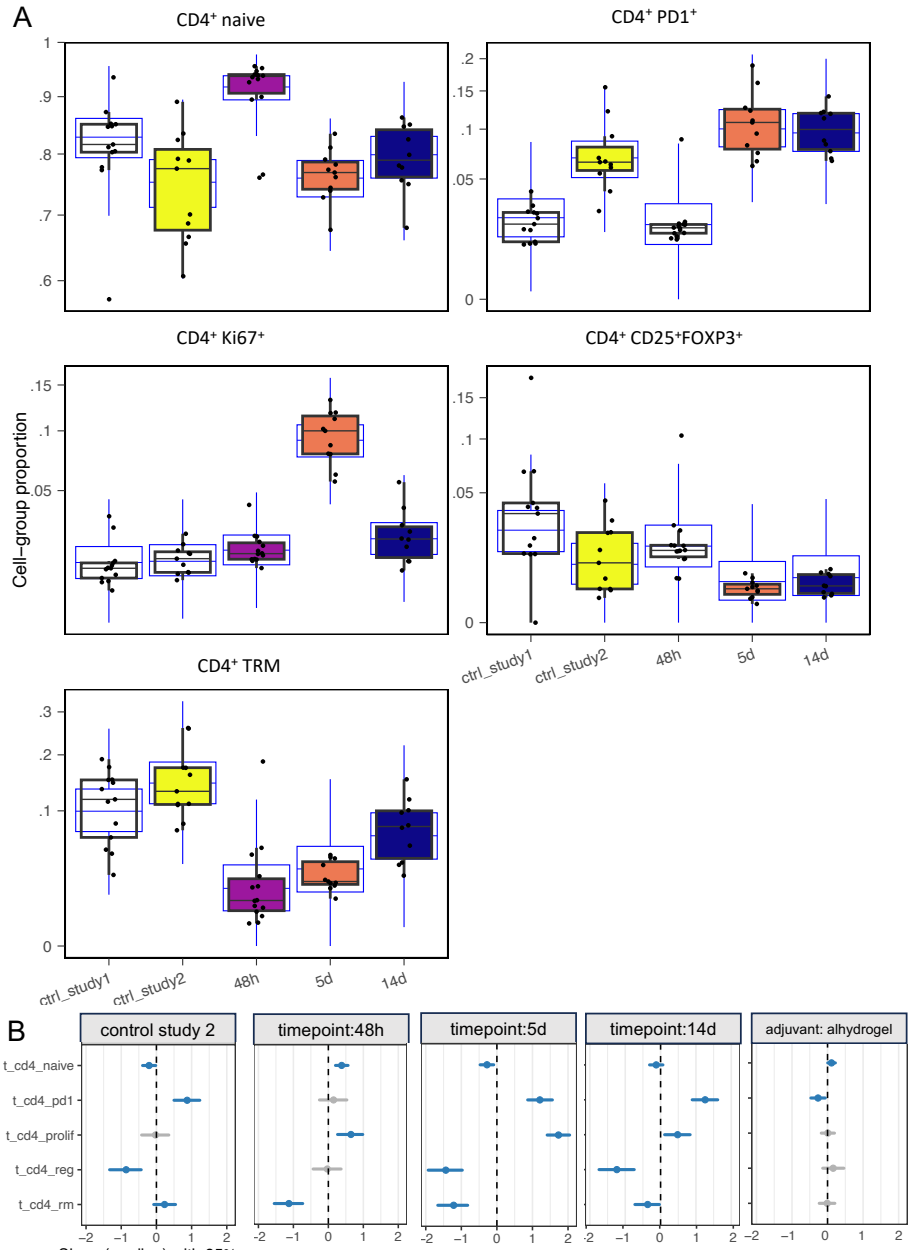

Figure 4.21: Compositional modelling results showing the temporal evolution of $CD4^+$ T cell phenotypes following intradermal KLH rechallenge. Results are aggregated across the KLH1 and KLH2 studies. $CD4^+$ T cell phenotypic proportions are modelled within each condition and timepoint using Bayesian compositional models. Statistical significance is assessed by comparison with the KLH1 control condition (ctrl_study1), which serves as the reference category for hypothesis testing.
(A) Boxplots summarising observed proportions of $CD4^+$ T cell phenotypes at each timepoint. Boxplots are filled when the corresponding timepoint differs significantly from the KLH1 control at a 5% FDR, with distinct fill colours indicating different timepoints. (B) Interval plots showing posterior mean estimates and 95% credible intervals for the difference in expected $CD4^+$ T cell phenotype proportion relative to the KLH1 control (logit scale) at each timepoint. Intervals are coloured blue when the difference is statistically significant at a 5% FDR and grey otherwise.

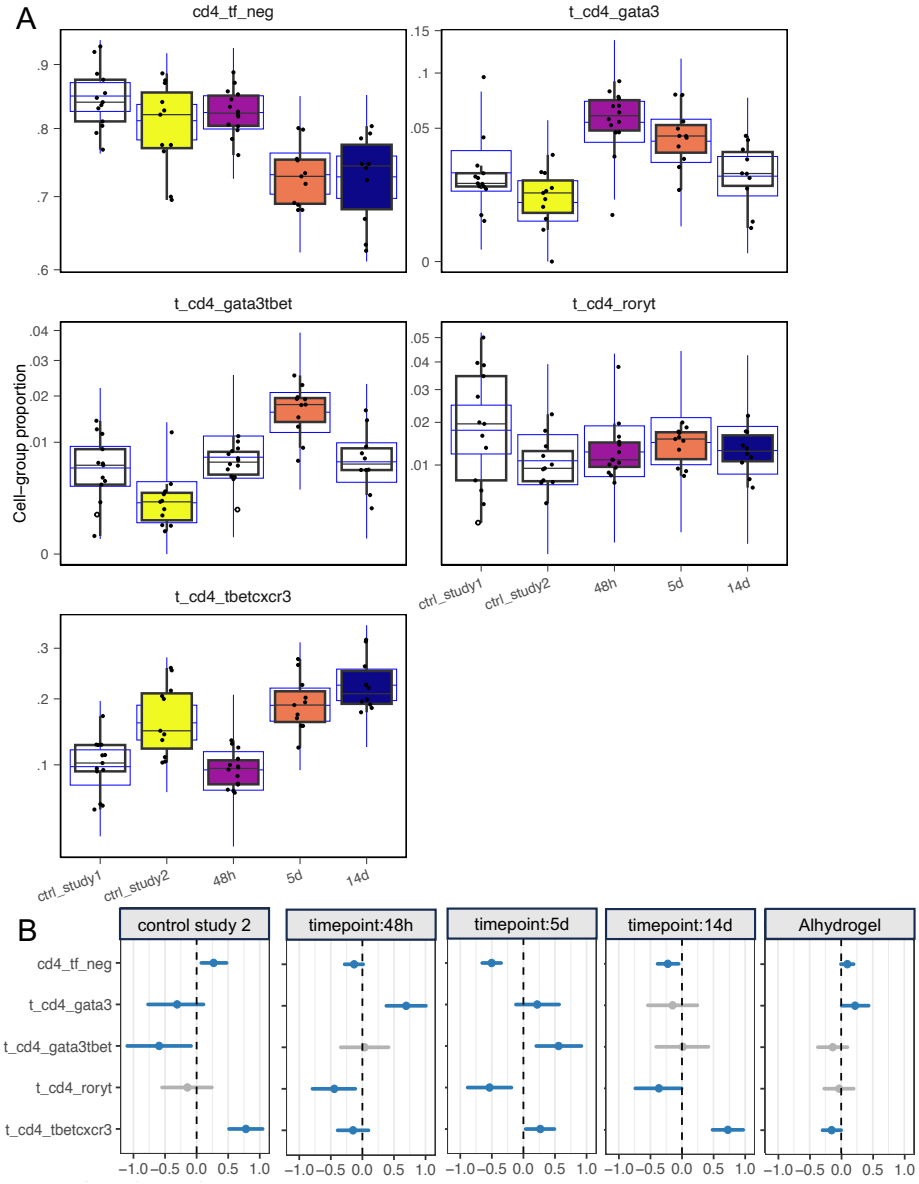

Figure 4.22: Compositional modelling results showing the temporal evolution of $CD4^+$ T cell subsets defined by transcription factor expression following intradermal KLH rechallenge. Results are aggregated across the KLH1 and KLH2 studies. $CD4^+$ T cell subset proportions are modelled within each condition and timepoint using Bayesian compositional models. Statistical significance is assessed by comparison with the KLH1 control condition (ctrl_study1), which serves as the reference category for hypothesis testing. Transcription factor–defined subsets include T-bet^+CXCR3^+ (Th1), $GATA3^+$ (Th2), and $ROR\gamma t^+$ (Th17) cells. Cells exhibiting mixed transcription factor expression (T-bet^+GATA3^+) and transcription factor–negative cells are also shown.
(A) Boxplots summarising observed subset proportions at each timepoint. Boxplots are filled when the corresponding timepoint differs significantly from the KLH1 control at a 5% FDR, with distinct fill colours indicating different timepoints. (B) Interval plots showing posterior mean estimates and 95% credible intervals for the difference in expected subset proportion relative to the KLH1 control (logit scale) at each timepoint. Intervals are coloured blue when the difference is statistically significant at a 5% FDR and grey otherwise.

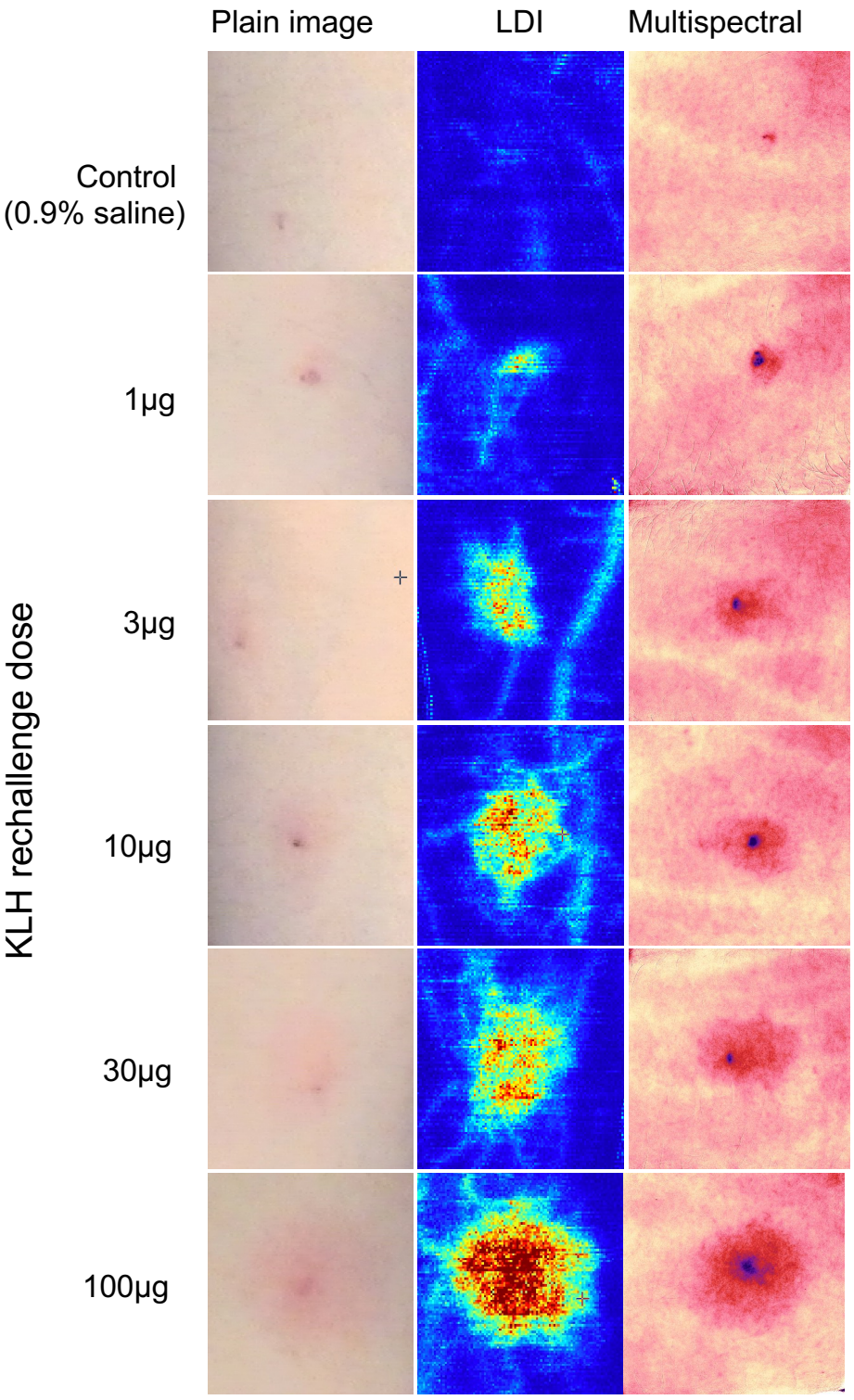

Figure 4.23: Illustrative comparison of dose-response in skin 48 h post-KLH rechallenge, for a representative participant in the KLH1 study. Left: plain image, middle: LDI, right: Antera Multispectral imaging.

rics quantifying blood flow (LDI mean flux, Antera Redness) were positively correlated albeit less strongly (Figure 4.24).

Prediction of cellular infiltrate We next benchmarked these metrics against the ground truth of the cellular infiltrate (flow cytometry of punch biopsies). Clinical induration demonstrated the strongest correlation with the absolute magnitude of both the lymphoid ($CD4^+$ T cells) and myeloid (total monocytes) infiltrate (Figure 4.25). In contrast, imaging metrics—which primarily quantify vascular dilation and perfusion—showed weaker associations with cellular abundance. Thus clinical measurement of oedema appears to be the non-invasive measure of choice, with other modalities playing an adjunctive role.

4.3.10 Participant experience and acceptability

To assess the feasibility and acceptability of the KLH challenge model, we administered a feedback survey to all participants in the KLH1 cohort using Visual Analogue Scales (VAS).

Motivation for participation Participants reported various motivations for enrolling in the study. Helping others and aiding scientific discovery were consistently ranked as highly important drivers, alongside personal interest in the research topic and financial compensation (Figure 4.26).

Tolerability of procedures The study procedures were generally well-tolerated. Participants reported mild to moderate discomfort, which was primarily attributable to the immediate transient sensation of the intradermal injections for rechallenge and during skin anaesthesia for punch skin biopsy (Figure 4.27).

Overall acceptability Overall satisfaction with the study experience was high. The majority of participants indicated a strong likelihood of considering similar studies in the future and stated they would recommend participation to friends or family members (Figure 4.28). These data support the viability of the KLH challenge framework as a sustainable experimental platform.

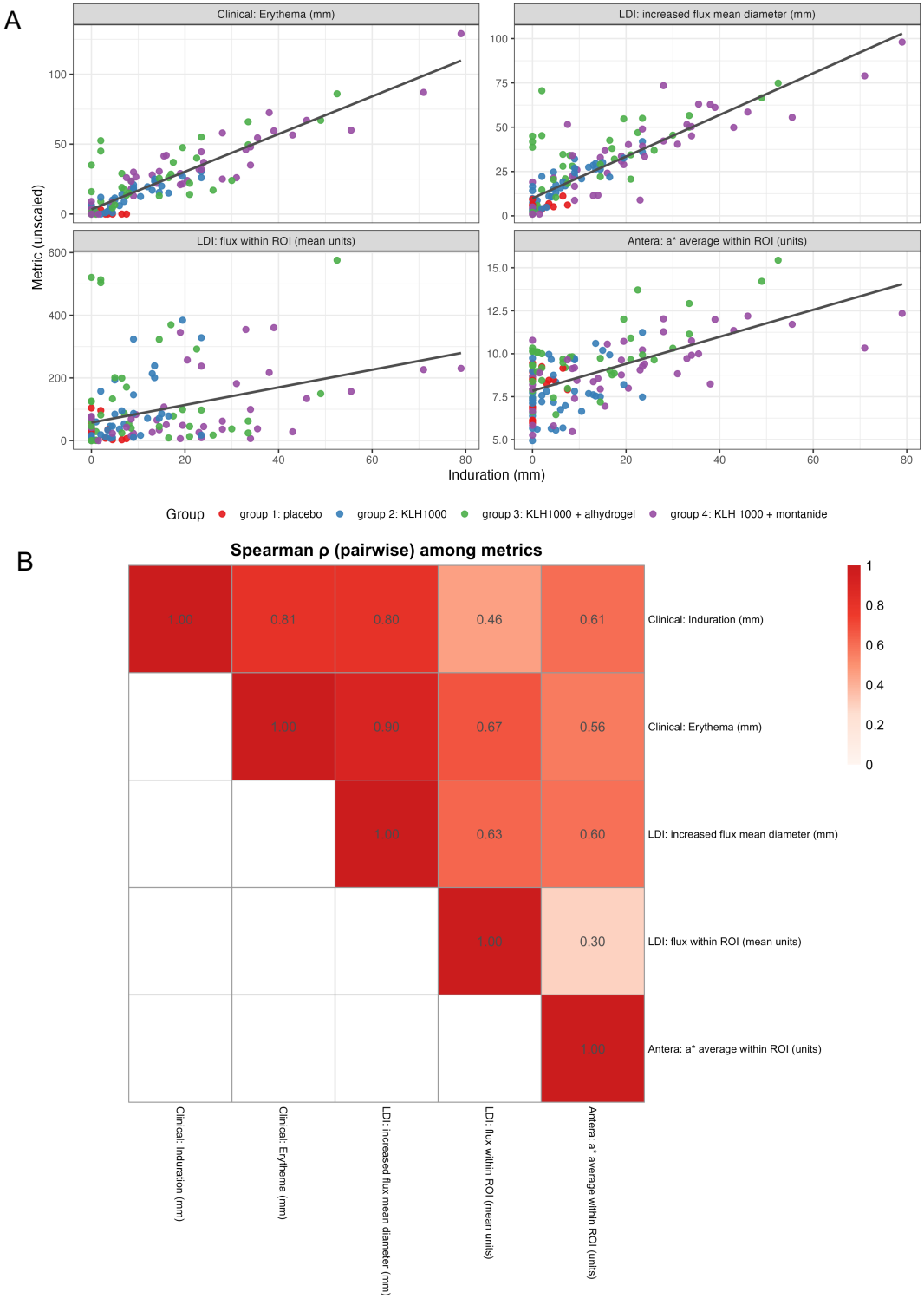

Figure 4.24: Agreement between clinical assessment and non-invasive imaging metrics for cutaneous KLH rechallenge responses. Data pooled across rechallenge dose levels (KLH1 study)
(A) Scatter plots comparing clinical induration diameter (reference standard) with clinical erythema and objective non-invasive imaging metrics derived from laser Doppler imaging (LDI) and multispectral imaging (Antera). (B) Heatmap of pairwise Spearman rank correlation coefficients (ρ) summarising monotonic associations between all clinical and imaging-derived metrics.

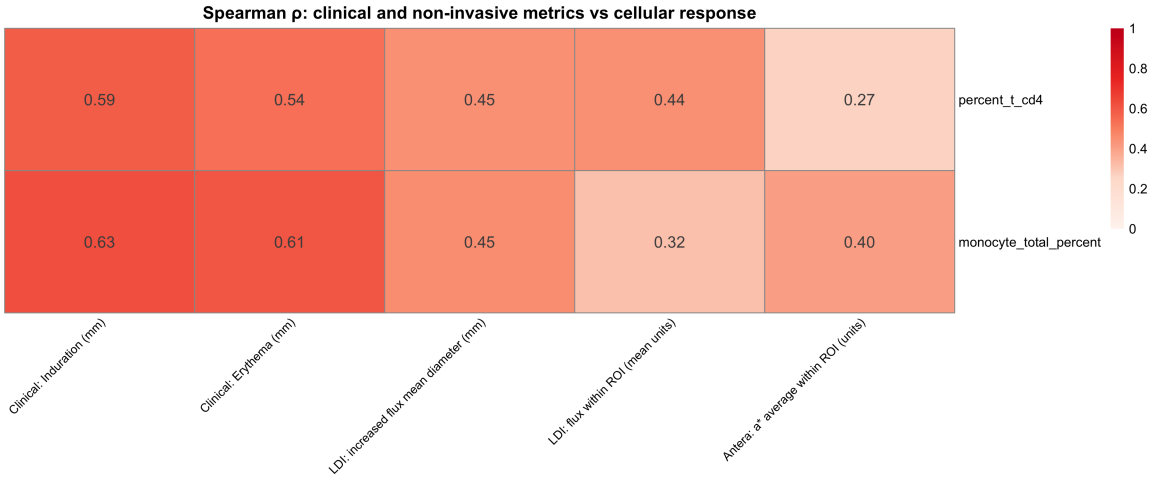

Figure 4.25: Correlation of non-invasive clinical and imaging-derived metrics with the underlying cellular infiltrate at active KLH rechallenge sites.
The heatmap displays Spearman rank correlation coefficients (ρ) between non-invasive predictors (columns) and flow cytometric measures of immune cell abundance (rows), including $CD4^+$ T cells and total monocytes. This analysis benchmarks the ability of non-invasive metrics to reflect the magnitude of the cellular response within the skin.

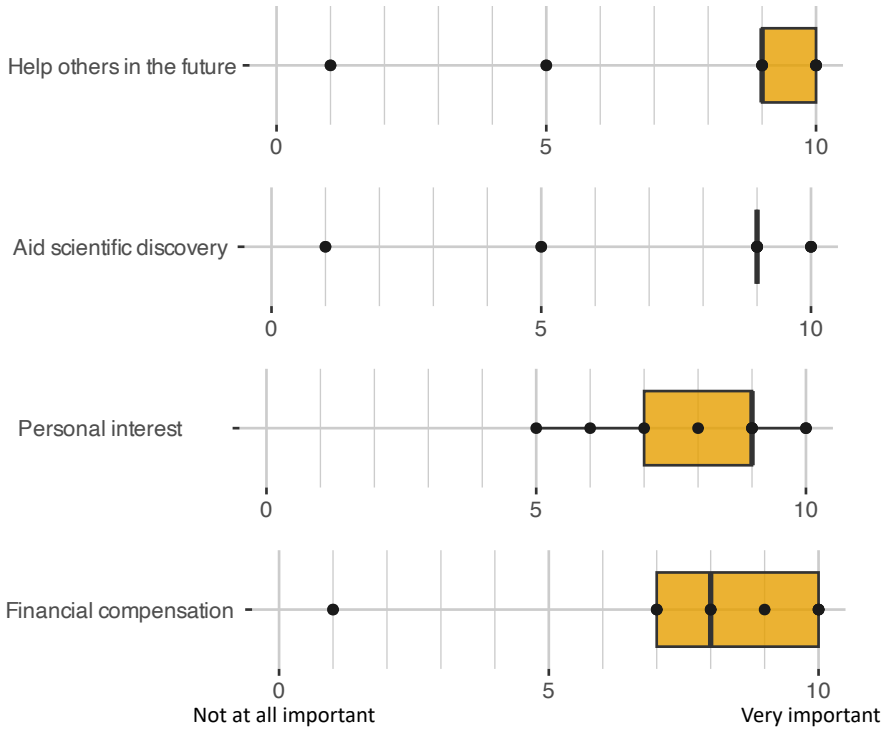

Figure 4.26: Self-reported motivations for study participation. Responses were recorded using a visual analogue scale (VAS) ranging from 0 (not important) to 10 (very important).

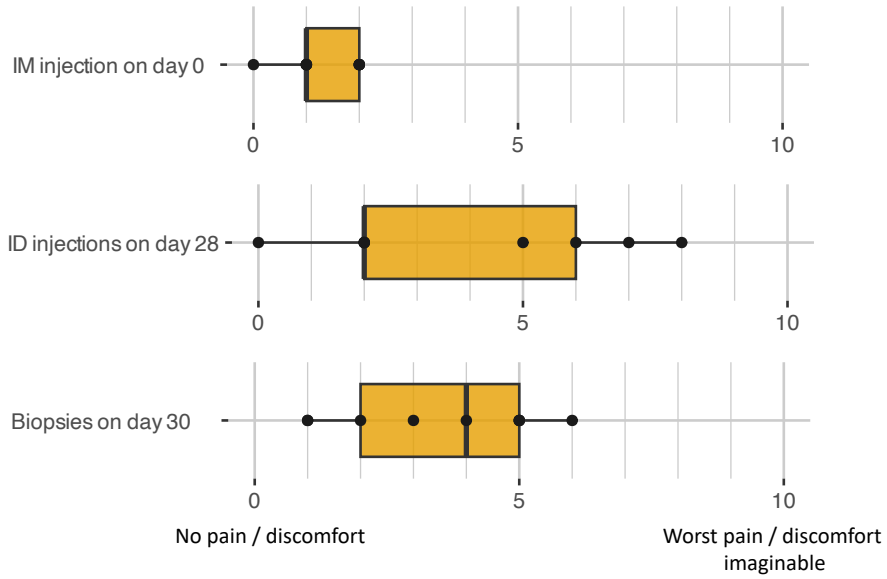

Figure 4.27: Participant-reported discomfort associated with study procedures. Discomfort was rated using a visual analogue scale (VAS) ranging from 0 (no discomfort) to 10 (worst imaginable discomfort). Intradermal injection corresponds to the peak pain experienced during administration of the rechallenge dose.

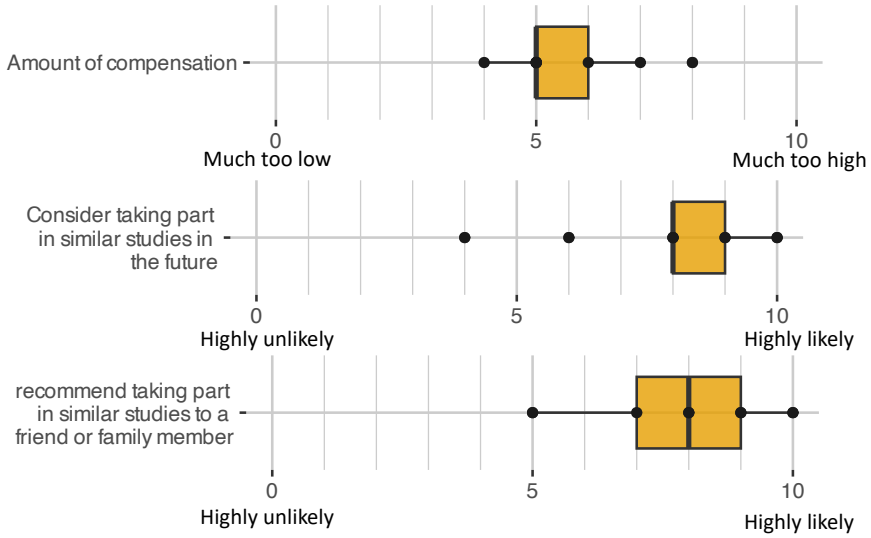

Figure 4.28: Overall study acceptability. Participants rated their likelihood of participating in future studies or recommending the study to others using a visual analogue scale (VAS) ranging from 0 (highly unlikely) to 10 (highly likely).

4.4 Discussion

4.4.1 Principal findings

In this study, we systematically optimised the KLH HIC model for use in future proof-of-mechanism studies. We established that adjuvant selection is a critical determinant of systemic immunogenicity: Montanide generated robust and consistent systemic T cell responses (IFNγ and IL-4 ELISpot) that were superior to Alhydrogel, although both adjuvants elicited comparable high-titre IgG responses. In contrast, unadjuvanted KLH failed to generate a reliable systemic cellular or humoral response. Despite this lack of systemic priming, unadjuvanted KLH was sufficient to sensitise participants for a detectable, albeit smaller, cutaneous response upon intradermal rechallenge. This dissociation suggests that the intradermal challenge model may be more sensitive for detecting low-level sensitisation than peripheral blood sampling. Through Bayesian Emax modelling of the cutaneous response, we identified 10 μg as the optimal rechallenge dose for these primary regimens, balancing detectability

and tolerability. Finally, detailed temporal profiling revealed a dynamic evolution of the local immune response. The classic 48 h observation window represents a mixed classical monocyte / naive T cellinfiltrate with balanced Th1/Th2 features. In contrast, Day 5 and Day 14 biopsies yielded a purer T cell infiltrate with a progressive accumulation of and T resident memory cells consistent with local tissue adaptation.

4.4.2 Adjuvant selection and immune polarisation

In addition to use of adjuvants to enhance the magnitude of the resultant immune response to KLH, we hypothesised that Alhydrogel would drive a Th2 profile and Montanide a mixed Th1/Th2 profile consistent with their known mechanisms-of-action. While Montanide successfully drove a potent mixed response (high IFNγ and IL-4), the "Th2-skewing" effect of Alhydrogel was less distinct in this model, manifesting primarily as a weaker overall cellular response rather than a pure Th2 phenotype. For general PoM studies where the goal is maximizing the signal-to-noise ratio of the pharmacodynamic biomarker, Montanide may be the preferred choice due to the magnitude and consistency of the elicited response.

4.4.3 Optimising the challenge: Dose and Time

A key finding of this work is the statistical definition of the challenge parameters. The dose-response modelling supports 10 μg as a standard rechallenge dose; it sits towards the top of the linear portion of the dose-response curve for sensitised individuals, ensuring sensitivity to modulation by IMPs, while avoiding the saturation and potential adverse events associated with high dose levels.[91] This operating characteristic is particularly important for PoM studies, where excessive responses may obscure meaningful PD effects. This finding noted, other doses may be justified depending on the specific objectives of the study. It should be noted that the rechallenge dose must be interpreted in the context of the primary immunisation regimen. For example, it is notable that while, in this study, detectable induration was observed at the lower dose level of 1 μg, other studies which used alternative regimens (e.g. 100μg KLH + Alhydrogel x3 doses at 2-weekly intervals) failed to yield

any detectable induration.[94]

Equally important is the selection of the sampling timepoint. Our data suggests that the optimal timepoint likely depends on the mechanism of action of the IMP. Drugs targeting myeloid biology or early inflammatory cytokines (e.g., anti-TNF, anti-IL6) may be best assessed at 48 h. Conversely, drugs targeting local T cell phenotypic adaptation, Th1 polarisation, or B cell/T cell interactions may be better assessed at Day 5 or Day 14, especially for bulk-RNA seq based endpoints where the transcriptomic signature represents an admixture of multiple cell types, thus suggesting a preference for later time points where the myeloid infiltrate has largely resolved.

4.4.4 Outcomes assessment: Imaging vs Biopsy

We evaluated non-invasive imaging (LDI, Multispectral) alongside clinical and histological endpoints. While imaging metrics correlated well with clinical induration and provide a useful, objective audit trail or inter-observer standardisation tool (especially for multi-centre studies), they do not serve as a surrogate for biopsy and do not appear to offer clear advantages over the clinical assessment of induration. The rich immunological resolution provided by flow cytometry and transcriptomics—especially immune processes occurring beyond the macroscopic clinical peak—cannot be captured by the imaging modalities tested. It is possible that other advanced high-resolution non-invasive modalities, such as reflectance confocal microscopy are capable of capturing cellular dynamics missed by less sensitive tools—this should be evaluated in future studies.

4.4.5 Limitations

Several limitations should be noted. Firstly, the sample sizes of both KLH1 and KLH2 studies are small, limiting the power to detect all but strong effects such as the effects of adjuvants primary response, and the effect of dose on rechallenge response. Subtle or moderate effects, such as the effect of primary adjuvant on the transcriptomic response to rechallenge, may have been missed in this context. The small sample size also results in broad uncertainty intervals for key immune parameters. Our transcriptomic analysis was limited by sample availability, preventing a direct comparison of

active responses to control tissue. The transcriptomic results should be therefore regarded as pilot data. Indeed, these findings prompted development of enhanced tissue processing protocols which are anticipated to result in substantially improved RNA yields in future HIC studies.

4.5 Conclusions and future directions

This study has mapped critical determinants of the response to KLH and the timecourse of rechallenge responses, and therefore represents a significant step towards a fit-for-purpose KLH HIC paradigm. A critical next step is to perform larger validation studies focusing on specific assays, endpoints, and timepoints of assessment to establish the statistical operating characteristics of key PD biomarker candidates. Subsequently, it would be of great value to perform benchmarking studies using immunomodulatory drugs with well-understood mechanisms and clinical efficacy. Such studies, performed systematically, would provide insights into the magnitude of modulation which may be seen for specific biomarkers, thereby assisting with design assumptions for IMPs targeting similar pathways.

4.6 Supplementary Methods and Results

Table 4.4: Flow cytometry panel reagents used for skin biopsy characterisation

Target Specificity	**Clone**	**Fluorochrome**
TIM-4	921832R	AF350
α-Smooth Muscle Actin	EPR5368	Alexa Fluor 594
S100A4	1F11	Alexa Fluor 700
CD90 (THY1)	eBio5E10 [5E10]	APC-eFluor 780
CD14	63D3	APC-Fire 810
CD28	CD28.2	BB515
CD25	M-A251	BB700

Continued on next page

Table 4.4 – continued from previous page

Target Specificity	Clone	Fluorochrome
CD27	O323	BUV496
MHC Class II (HLA-DR, DQ)	L243	BUV563
CD206	19.2	BUV615
CD38	HIT2	BUV661
CD103	Ber-ACT8	BUV737
CD3	UCHT1	BUV805
TCR $\alpha\beta$	IP26	BUV805
IgD	IA6-2	BV480
CD4	SK3	BV510
CD45	HI30	BV570
CD69	FN50	BV605
CD141	1A4	BV650
CD10	HI10a	BV750
CD163	GHI/61	BV786
CD19	HIB19	cFluor B532
CD20	2H7	cFluor B675
CD56	LT56	cFluor BYG750
CD16	eBioCB16 [CB16]	eFluor 450
CD279 (PD-1)	MIH4	PE-CF594
CD123	9F5	PE-Cy5
MERTK	590H11G1E3	PE-Cy7
CD127	A019D5	PE-Fire 640
CD183 (CXCR3)	G025H7	PE-Fire 810
CD1c	L161	PerCP-Cy5.5
CD8	SK1	Spark Blue 550
CD45RA	HI100	Spark YG 581
Viability	-	Zombie NIR

Continued on next page

Table 4.4 – continued from previous page

Target Specificity	Clone	Fluorochrome
Intracellular / Transcription Factors		
T-bet	4B10	Alexa Fluor 647
GATA3	L50-823	BUV395
FOXP3	206D	BV421
Ki-67	Ki-67	BV711
RORγ(t)	Q21-559	PE
IRF4	3E4	PerCP-eFluor 710

4.6.1 Additional description of RNAseq methods

Alignment and QC were performed by Dr Patrick MacLean. The pipeline was executed with Nextflow v24.04.2 (Di Tommaso et al., 2017) with the following command:

```
nextflow run [path-to-data]] \
--outdir [path-to-output] \
--input [path-to-data]] \
-c []path-to-config] --skip_biotype_qc \
--aligner star_salmon --cleanup true \
--email [email-address] \
--email_on_fail [email-address] \
--genome GRCh38
```

The following software versions were used:

Process Name	Software	Version
BEDTOOLS_GENOMECOV	bedtools	2.30.0
CUSTOM_DUMPSOFTWAREVERS	python	3.11.7
	yaml	5.4.1
CUSTOM_GETCHROMSIZES	getchromsizes	1.16.1
DESEQ2_QC_STAR_SALMON	bioconductor-deseq2	1.28.0
	r-base	4.0.3
DUPRADAR	bioconductor-dupradar	1.28.0
	r-base	4.2.1
FASTQC	fastqc	0.12.1
FQ_SUBSAMPLE	fq	0.9.1 (2022-02-22)
GTF_FILTER	python	3.9.5
MAKE_TRANSCRIPTS_FASTA	rsem	1.3.1
	star	2.7.10a
PICARD_MARKDUPLICATES	picard	3.0.0
QUALIMAP_RNASEQ	qualimap	2.3
RSEQC_BAMSTAT	rseqc	5.0.2
RSEQC_INFEREXPERIMENT	rseqc	5.0.2
RSEQC_INNERDISTANCE	rseqc	5.0.2
RSEQC_JUNCTIONANNOTATI(	rseqc	5.0.2
RSEQC_JUNCTIONSATURATI(	rseqc	5.0.2
RSEQC_READDISTRIBUTION	rseqc	5.0.2
RSEQC_READDUPLICATION	rseqc	5.0.2
SALMON_INDEX	salmon	1.10.1
SALMON_QUANT	salmon	1.10.1
SAMTOOLS_FLAGSTAT	samtools	1.17
SAMTOOLS_IDXSTATS	samtools	1.17
SAMTOOLS_INDEX	samtools	1.17
SAMTOOLS_SORT	samtools	1.17
SAMTOOLS_STATS	samtools	1.17

(continued on next page)

Process Name	Software	Version
SE_GENE	bioconductor-summarizedexperiment	1.24.0
	r-base	4.1.1
STAR_ALIGN_IGENOMES	gawk	5.1.0
	samtools	1.10
	star	2.6.1d
STRINGTIE_STRINGTIE	stringtie	2.2.1
TRIMGALORE	cutadapt	3.4
	trimgalore	0.6.7
TX2GENE	python	3.9.5
TXIMPORT	bioconductor-tximeta	1.12.0
	r-base	4.1.1
UCSC_BEDCLIP	ucsc	377
UCSC_BEDGRAPHTOBIGWIG	ucsc	445
Workflow	Nextflow	24.04.2
	nf-core/rnaseq	3.14.0

4.6.2 Supplementary tables

Table 4.6: Baseline laboratory values by randomisation group)

Laboratory parameter	N	Randomisation group				
		Placebo N = 3	**KLH 1000** N = 7	**KLH 1000 + Alhydrogel** N = 7	**KLH 1000 + Montanide** N = 7	**Overall** N = 24
Haemoglobin (g/L)	24	128.0 [125.0, 147.0]	136.0 [129.0, 145.0]	133.0 [127.0, 157.0]	148.0 [137.0, 156.0]	138.0 [128.5, 150.5]
WCC (×10^9/L)	24	5.8 [5.1, 6.5]	5.0 [4.2, 6.0]	6.0 [4.9, 6.8]	6.8 [4.4, 7.5]	5.7 [4.6, 6.8]
Platelets (×10^9/L)	24	213.0 [205.0, 239.0]	246.0 [202.0, 311.0]	254.0 [250.0, 265.0]	243.0 [236.0, 264.0]	248.0 [222.5, 264.5]
Neutrophils (×10^9/L)	24	3.0 [3.0, 3.5]	2.5 [2.3, 3.3]	3.4 [3.1, 3.7]	3.6 [2.7, 5.2]	3.1 [2.5, 3.6]
Lymphocytes (×10^9/L)	24	2.0 [1.4, 2.1]	1.9 [1.2, 2.1]	2.1 [1.2, 2.4]	1.6 [1.2, 2.5]	1.9 [1.3, 2.1]
Monocytes (×10^9/L)	24	0.6 [0.5, 0.6]	0.4 [0.3, 0.6]	0.4 [0.4, 0.5]	0.5 [0.3, 0.7]	0.5 [0.4, 0.5]
Sodium (mmol/L)	23	140.0 [137.0, 142.0]	141.0 [140.0, 142.0]	140.5 [140.0, 142.0]	141.0 [138.0, 143.0]	141.0 [140.0, 142.0]
Unknown		0	0	1	0	1
Potassium (mmol/L)	24	4.3 [4.1, 4.6]	4.0 [3.9, 4.3]	4.1 [3.7, 4.6]	4.2 [4.1, 4.7]	4.2 [4.0, 4.5]
Urea (mmol/L)	23	4.1 [3.0, 5.6]	4.4 [3.9, 4.8]	4.4 [3.4, 4.8]	5.3 [4.0, 6.5]	4.4 [4.0, 5.6]
Unknown		0	0	1	0	1
Creatinine (µmol/L)	24	67.0 [53.0, 69.0]	68.0 [55.0, 80.0]	57.0 [54.0, 82.0]	79.0 [68.0, 80.0]	68.5 [56.5, 80.0]

[1] Median [Q1, Q3]

Continuous variables are median [IQR].

Table 4.7: Serology (anti-KLH IgG) on $\log_{10}$ scale: arithmetic mean (95% CI)

Randomisation group	Time	N	Estimate (95% CI)
Placebo	D0	3	4.08 (3.66–4.49)
Placebo	D7	3	3.87 (2.52–5.22)
Placebo	D14	3	3.94 (2.42–5.45)
Placebo	D21	3	4.16 (3.7–4.62)
Placebo	D28	3	4.22 (3.9–4.55)
KLH 1000	D0	7	4.19 (4–4.38)
KLH 1000	D7	7	4.28 (4.06–4.49)
KLH 1000	D14	7	4.4 (4.14–4.66)
KLH 1000	D21	7	4.34 (4.15–4.52)
KLH 1000	D28	7	4.29 (4.14–4.45)
KLH 1000 + Alhydrogel	D0	7	4.34 (4.09–4.59)
KLH 1000 + Alhydrogel	D7	7	4.22 (3.87–4.56)
KLH 1000 + Alhydrogel	D14	7	5.2 (4.81–5.58)
KLH 1000 + Alhydrogel	D21	7	5.35 (4.91–5.79)
KLH 1000 + Alhydrogel	D28	7	5.33 (4.9–5.77)
KLH 1000 + Montanide	D0	7	4.21 (3.96–4.47)
KLH 1000 + Montanide	D7	7	4.17 (3.86–4.48)
KLH 1000 + Montanide	D14	7	5.04 (4.76–5.32)
KLH 1000 + Montanide	D21	7	5.28 (4.96–5.61)
KLH 1000 + Montanide	D28	7	5.39 (4.99–5.8)

Values are $\log_{10}$-transformed; arithmetic mean with 95% CI (t-interval). N = non-missing at time-point/arm.

4.6.3 Diagnostics for the Emax model

Review of the cmdstanr diagnostics suggested satisfactory tree depth for all transition, and no evidence of divergent transitions. E-BFMI was satisfactory, as were the effective sample size and R-hat values.

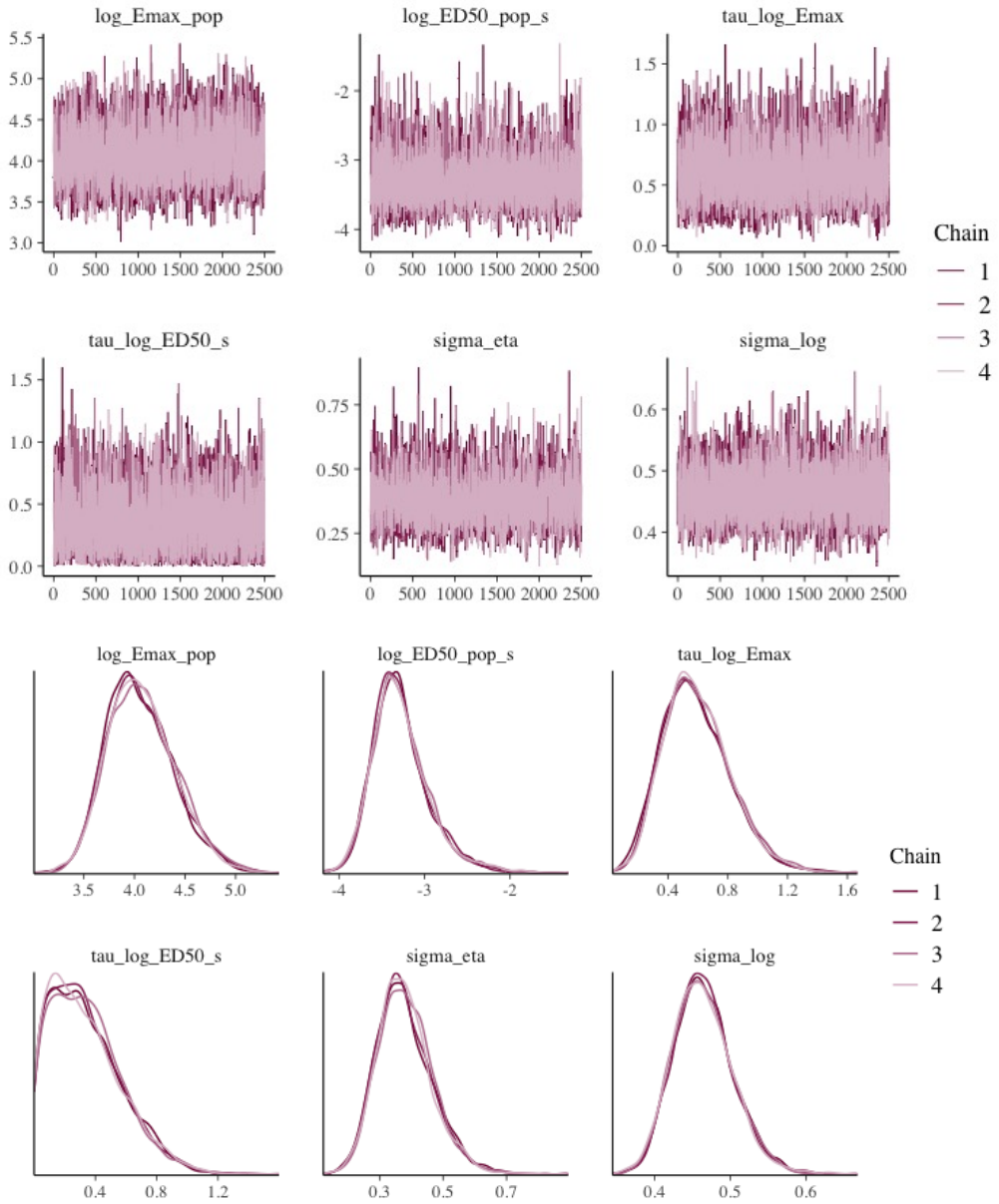

Figure 4.29: Hamiltonian Monte Carlo (HMC) diagnostics for the Bayesian E_{max} dose–response model. Panels show trace and marginal density plots for key parameters across four independent chains, demonstrating satisfactory mixing and convergence.

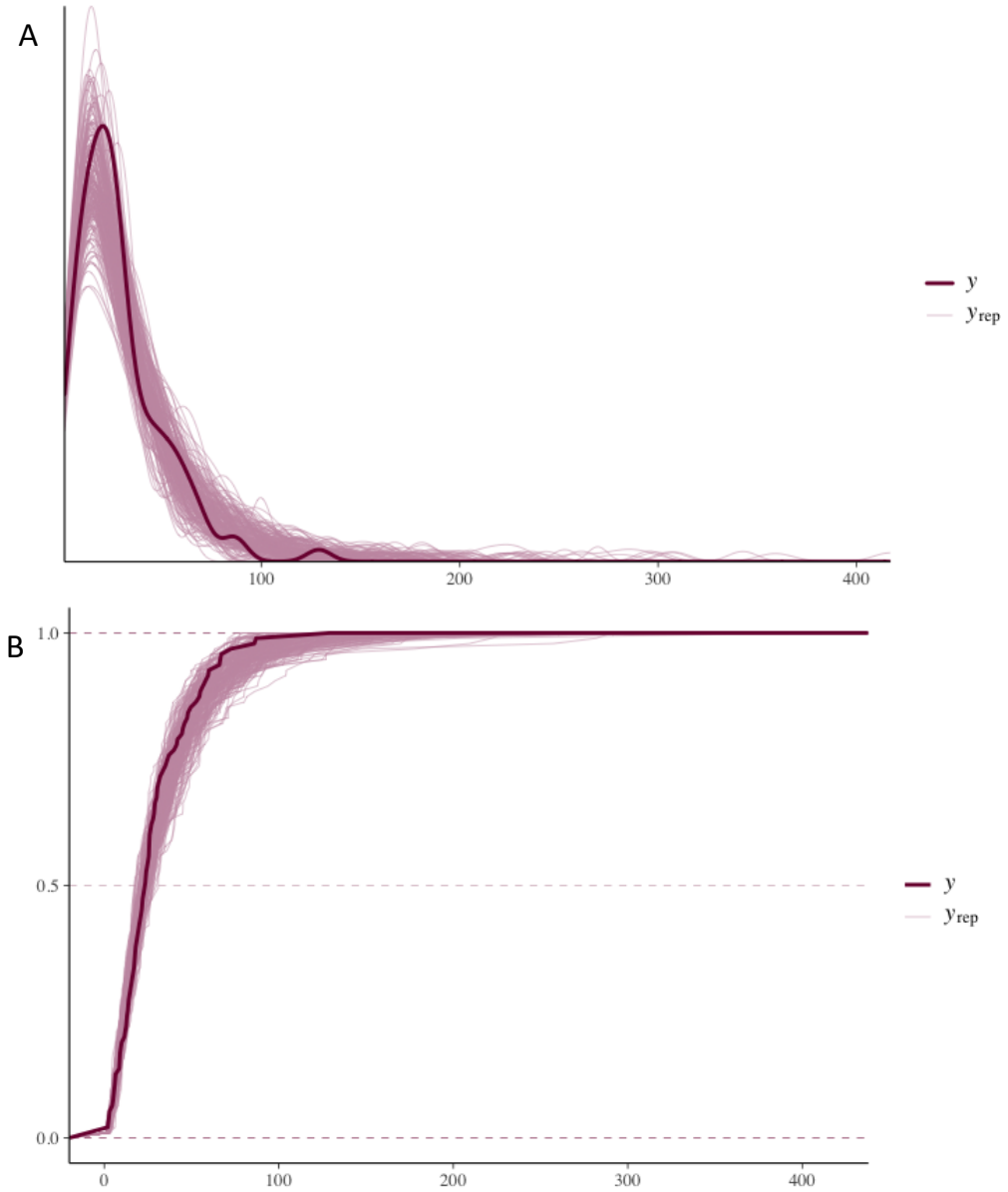

Figure 4.30: Posterior predictive checks (PPCs) assessing the fit of the Bayesian E_{max} dose–response model.
(A) Density overlay comparing the kernel density estimate of the observed induration data (y, dark line) with density estimates from 200 replicated datasets (y_{rep}, light red lines) drawn from the posterior predictive distribution. (B) Empirical cumulative distribution function (ECDF) overlay comparing the cumulative distribution of the observed data with replicated datasets.
In both panels, the observed data lie centrally within the envelope of the simulated data, indicating that the model adequately captures the central tendency, dispersion, and tail behaviour of the observed clinical response.

Table 4.8: ELISpot IFNγ: geometric mean (95% CI)

Randomisation group	Time	N	Estimate (95% CI)
Placebo	D0	2	19.4 (NA)
Placebo	D7	2	27.5 (NA)
Placebo	D14	2	42.4 (NA)
Placebo	D21	2	7.32 (NA)
Placebo	D28	2	36.1 (NA)
KLH 1000	D0	5	14 (0–50.4)
KLH 1000	D7	5	5.88 (0.0565–49.8)
KLH 1000	D14	5	23.2 (0–83)
KLH 1000	D21	5	18.9 (0.399–155)
KLH 1000	D28	5	19.2 (8.03–46)
KLH 1000 + Alhydrogel	D0	6	15.6 (2.52–58.1)
KLH 1000 + Alhydrogel	D7	6	129 (16.6–1010)
KLH 1000 + Alhydrogel	D14	6	124 (18.3–559)
KLH 1000 + Alhydrogel	D21	6	216 (77–606)
KLH 1000 + Alhydrogel	D28	6	124 (34.8–442)
KLH 1000 + Montanide	D0	5	25.6 (2.56–90.8)
KLH 1000 + Montanide	D7	5	97.4 (0–1110)
KLH 1000 + Montanide	D14	5	2010 (962–4190)
KLH 1000 + Montanide	D21	5	1400 (595–3280)
KLH 1000 + Montanide	D28	5	1030 (519–2030)

Raw scale; geometric mean with 95% CI via log-transform. N = non-missing at timepoint/arm.

Table 4.9: ELISpot IL-4: geometric mean (95% CI)

Randomisation group	Time	N	Estimate (95% CI)
Placebo	D0	2	10.5 (NA)
Placebo	D7	2	4.27 (NA)
Placebo	D14	2	18.7 (NA)
Placebo	D21	2	2.44 (NA)
Placebo	D28	2	25.2 (NA)
KLH 1000	D0	5	3.48 (0–16.2)
KLH 1000	D7	5	8.68 (3.15–23.9)
KLH 1000	D14	5	36.6 (7.99–168)
KLH 1000	D21	5	38.8 (4.71–125)
KLH 1000	D28	5	41.7 (11.8–147)
KLH 1000 + Alhydrogel	D0	6	3.18 (0–25.9)
KLH 1000 + Alhydrogel	D7	6	76.3 (12.5–464)
KLH 1000 + Alhydrogel	D14	6	181 (32.4–1010)
KLH 1000 + Alhydrogel	D21	6	170 (55.7–516)
KLH 1000 + Alhydrogel	D28	6	169 (54.3–524)
KLH 1000 + Montanide	D0	5	5.35 (0–36.5)
KLH 1000 + Montanide	D7	5	75.6 (3–416)
KLH 1000 + Montanide	D14	5	957 (514–1780)
KLH 1000 + Montanide	D21	5	630 (267–1490)
KLH 1000 + Montanide	D28	5	467 (217–1010)

Raw scale; geometric mean with 95% CI via log-transform. N = non-missing at timepoint/arm.

Table 4.10: Serology (anti-KLH IgG, $\log_{10}$): $AUC_{0–28}$ arm contrasts (ANOVA)

Contrast	Estimate	95% CI	p (Holm)
Alhydrogel - KLH 1000	16.4	7.9 to 24.9	0.000
Montanide - KLH 1000	14.3	5.79 to 22.8	0.001
Montanide - Alhydrogel	-2.11	-10.6 to 6.39	0.521

Serology uses $\log_{10}$ values at D0,7,14,21,28; $AUC_{0text--28}$ via trapezoidal rule (days). One-way ANOVA; estimates are mean differences in $AUC_{0text--28}$ (rows shown as "B A"). Holm-adjusted p-values across the three planned contrasts. Placebo excluded.

Table 4.11: ELISpot IFNγ (frozen, raw): $AUC_{0–28}$ arm contrasts (ANOVA)

Contrast	Estimate	95% CI	p (Holm)
Alhydrogel - KLH 1000	6880	-16800 to 30600	0.440
Montanide - KLH 1000	34200	9390 to 59000	0.007
Montanide - Alhydrogel	27300	3560 to 51000	0.015

ELISpot IFNγ uses raw counts at D0,7,14,21,28; AUC_{0-28} via trapezoidal rule (days). One-way ANOVA; estimates are mean differences in AUC_{0-28}. Holm-adjusted p-values across the three planned contrasts. Placebo excluded.

Table 4.12: ELISpot IL-4 (frozen, raw): $AUC_{0–28}$ arm contrasts (ANOVA)

Contrast	Estimate	95% CI	p (Holm)
Alhydrogel - KLH 1000	5900	-3020 to 14800	0.093
Montanide - KLH 1000	14300	5020 to 23600	0.003
Montanide - Alhydrogel	8430	-483 to 17300	0.044

ELISpot IL-4 uses raw counts at D0,7,14,21,28; $AUC_{0text--28}$ via trapezoidal rule (days). One-way ANOVA; estimates are mean differences in $AUC_{0text--28}$. Holm-adjusted p-values across the three planned contrasts. Placebo excluded.

Table 4.13: Serology (anti-KLH IgG, $\log_{10}$ AUC) — ANCOVA ($AUC_{0–28}$ arm + age + sex + CMV)

Term	df	F	p	Partial eta^2
Study arm (Montanide vs Alhydrogel)	1	0.349	0.571	0.0418
Age (years)	1	0.0244	0.880	0.00304
Sex (Male vs Female)	1	0.423	0.534	0.0502
CMV (Detected vs Not detected)	1	0.00299	0.958	0.000374
Residuals	8	NA	NA	0.5

Alhydrogel and Montanide arms only. Complete cases (CMV 'Not performed' excluded). N = 13; residual df = 8; adjusted R^2 = -0.242.$Outcomeis\log_{10}$ transformed AUC (units: log_{10}-days).

Table 4.14: ELISpot IFNγ (raw AUC) — ANCOVA ($AUC_{0–28}$ arm + age + sex + CMV)

Term	df	F	p	Partial eta^2
Study arm (Montanide vs Alhydrogel)	1	4.43	0.089	0.47
Age (years)	1	0.318	0.597	0.0598
Sex (Male vs Female)	1	0.754	0.425	0.131
CMV (Detected vs Not detected)	1	2.82	0.154	0.361
Residuals	5	NA	NA	0.5

Alhydrogel and Montanide arms only. Complete cases (CMV 'Not performed' excluded). N = 10; residual df = 5; adjusted $R^2 = 0.515$. Outcome is raw AUC (units: spots·days).

Table 4.15: ELISpot IL-4 (raw AUC) — ANCOVA ($AUC_{0–28}$ arm + age + sex + CMV)

Term	df	F	p	Partial eta^2
Study arm (Montanide vs Alhydrogel)	1	4.63	0.084	0.481
Age (years)	1	2.62	0.167	0.344
Sex (Male vs Female)	1	0.0439	0.842	0.0087
CMV (Detected vs Not detected)	1	0.0875	0.779	0.0172
Residuals	5	NA	NA	0.5

Alhydrogel and Montanide arms only. Complete cases (CMV 'Not performed' excluded). N = 10; residual df = 5; adjusted $R^2 = 0.349$. Outcome is raw AUC (units: spots·days).

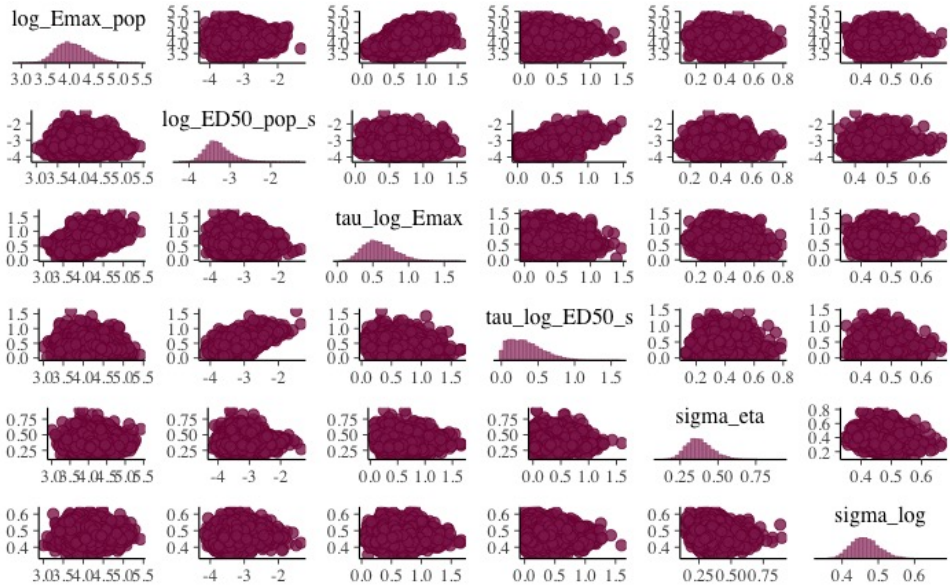

Figure 4.31: Pairs plot of key population-level parameters from the Bayesian E_{max} model, illustrating marginal posterior distributions and pairwise parameter relationships. The absence of strong non-linear dependencies or pathological correlations supports stable posterior geometry and reliable sampling.

Table 4.16: KLH dose–response summary (induration and erythema) by group and dose

Group	Dose	N	Induration mean (95% CI) [mm]	Erythema mean (95% CI) [mm]
Placebo	0	3	0.0 (0.0–0.0)	0.3 (0.0–1.0)
Placebo	1	3	2.5 (0.0–7.4)	0.8 (0.0–1.7)
Placebo	3	3	1.2 (0.0–3.5)	0.7 (0.0–1.3)
Placebo	10	3	1.5 (0.0–4.4)	0.3 (0.0–1.0)
Placebo	30	3	1.0 (0.0–3.0)	0.8 (0.0–1.7)
Placebo	100	3	2.8 (0.0–6.6)	1.7 (0.0–3.4)
KLH 1000	0	7	0.0 (0.0–0.0)	0.6 (0.0–1.2)
KLH 1000	1	7	2.5 (0.6–4.4)	4.9 (2.2–7.6)
KLH 1000	3	7	3.8 (0.6–7.0)	7.9 (2.9–12.9)
KLH 1000	10	7	7.9 (4.5–11.2)	12.7 (6.3–19.1)
KLH 1000	30	7	8.9 (4.8–13.1)	14.9 (8.8–21.0)
KLH 1000	100	7	15.4 (9.8–20.9)	18.4 (10.5–26.3)
KLH 1000 + Alhydrogel	0	7	0.0 (0.0–0.0)	0.3 (0.0–0.6)
KLH 1000 + Alhydrogel	1	7	6.5 (1.3–11.7)	17.1 (6.6–27.6)
KLH 1000 + Alhydrogel	3	7	8.3 (2.5–14.1)	13.3 (6.9–19.7)
KLH 1000 + Alhydrogel	10	7	13.8 (4.0–23.5)	21.1 (5.2–37.1)
KLH 1000 + Alhydrogel	30	7	13.9 (5.6–22.1)	32.4 (18.9–46.0)
KLH 1000 + Alhydrogel	100	7	25.4 (10.0–40.7)	46.2 (25.8–66.6)
KLH 1000 + Montanide	0	7	0.2 (0.0–0.6)	0.5 (0.0–1.5)
KLH 1000 + Montanide	1	7	15.1 (7.3–22.8)	23.3 (16.7–29.9)
KLH 1000 + Montanide	3	7	13.3 (8.5–18.1)	21.4 (15.6–27.2)
KLH 1000 + Montanide	10	7	23.1 (13.8–32.4)	35.8 (24.3–47.2)
KLH 1000 + Montanide	30	7	31.1 (16.5–45.8)	47.1 (33.4–60.9)
KLH 1000 + Montanide	100	7	45.4 (32.8–58.0)	70.0 (50.3–89.7)

4.6.4 Compositional model tables

The following tables present parameter estimates from Bayesian compositional models (sccomp) applied to flow cytometry data derived from enzymatically digested 5 mm punch biopsies collected 48 h after intradermal KLH rechallenge in the KLH1 study ($n = 21$). Each table corresponds to a distinct model as described in the Methods: Models 1 and 2 assess the association between clinical induration and cell-type composition; Model 3 assesses the effects of rechallenge dose and primary adjuvant on major immune cell populations; Models 4a and 4b examine these effects within the $CD4^+$ T cell compartment, focussing on classical phenotypes and transcription factor-defined subsets respectively. Parameter estimates are reported on the logit scale with posterior 95% credible intervals; significant associations are defined at a 5% FDR.

Table 4.17: Predicted composition by clinical response with 95% credible intervals (KLH1 model 1).

cell_group	negative	positive
B cell	0.0834% (0.052%–0.14%)	0.244% (0.184%–0.324%)
CD45-	79.2% (76.9%–81.4%)	43.7% (39.5%–47.9%)
Conventional DC	0.772% (0.575%–1.03%)	2.06% (1.63%–2.56%)
Macrophage	1.64% (1.26%–2.14%)	3.37% (2.54%–4.43%)
Classical monocyte	2.95% (2.19%–3.91%)	8.12% (6.54%–9.95%)
Intermediate monocyte	0.357% (0.277%–0.462%)	1.08% (0.797%–1.44%)
Non-classical monocyte	0.236% (0.183%–0.304%)	0.328% (0.254%–0.423%)
NK cell	0.0741% (0.051%–0.108%)	0.137% (0.104%–0.182%)
NK–T cell	0.0507% (0.0286%–0.0844%)	0.0565% (0.0419%–0.0761%)
pDC	0.111% (0.0823%–0.15%)	0.187% (0.146%–0.245%)
CD4+ T cell	10.2% (8.48%–12.3%)	31.6% (27.8%–35.6%)
CD8+ T cell	4.31% (3.34%–5.48%)	9.16% (7.58%– 11%)

Table 4.18: Predicted composition by dose with 95% credible intervals (KLH1 model 3).

cell_group	1	100
B cell	0.372% (0.248%–0.53%)	0.274% (0.17%–0.462%)
CD45-	31.2% (24%–39.2%)	39% (34.2%–43.3%)
Conventional DC	1.73% (1.26%–2.41%)	2.29% (1.84%–2.89%)
Macrophage	6.58% (4.76%–8.89%)	4.37% (3.21%– 5.9%)
Classical monocyte	9.43% (6.82%–12.9%)	11.3% (9.13%–14.1%)
Intermediate monocyte	1.21% (0.799%–1.78%)	2.01% (1.5%–2.83%)
Non-classical monocyte	0.345% (0.234%–0.519%)	0.515% (0.37%–0.752%)
NK cell	0.191% (0.123%–0.289%)	0.0958% (0.0622%–0.143%)
NK–T cell	0.129% (0.0861%–0.189%)	0.0289% (0.0204%–0.0408%)
pDC	0.246% (0.154%–0.381%)	0.145% (0.0927%–0.236%)
CD4+ T cell	36.3% (29.6%–43.4%)	31.1% (27.3%–35.5%)
CD8+ T cell	12.3% (8.76%–16.2%)	8.88% (5.94%–12.5%)

Table 4.19: Predicted composition by adjuvant with 95% credible intervals (KLH1 model 3).

cell_group	KLH_montanide	KLH_alhydrogel
B cell	0.369% (0.251%–0.533%)	0.319% (0.178%–0.626%)
CD45-	31.4% (24.3%–38.7%)	29.9% (20.3%–38.9%)
Conventional DC	1.76% (1.24%–2.47%)	2.21% (1.46%–3.33%)
Macrophage	6.64% (4.7%–9.18%)	3.68% (2.34%–5.57%)
Classical monocyte	9.43% (6.94%–12.8%)	7.7% (5.31%–11.1%)
Intermediate monocyte	1.23% (0.788%–1.87%)	0.912% (0.503%–1.69%)
Non-classical monocyte	0.351% (0.23%–0.519%)	0.306% (0.161%–0.582%)
NK cell	0.191% (0.124%–0.283%)	0.126% (0.0876%–0.177%)
NK–T cell	0.129% (0.0871%–0.192%)	0.0818% (0.056%–0.121%)

Table 4.19: Predicted composition by adjuvant with 95% credible intervals (KLH1 model 3). *(continued)*

cell_group	KLH_montanide	KLH_alhydrogel
pDC	0.245% (0.15%–0.39%)	0.258% (0.128%–0.501%)
CD4+ T cell	36.1% (29.5%–43.2%)	42.6% (33.1%–52.5%)
CD8+ T cell	12.2% (8.86%–16.3%)	11.9% (9.12%–15.3%)

Table 4.20: Predicted composition by dose with 95% credible intervals (KLH1 model 4a).

cell_group	1	100
CD4 naive	91.5% (88.7%–93.4%)	92.6% (91.6%–93.7%)
CD4 PD-1+	1.92% (1.39%–2.59%)	1.43% (1.18%–1.73%)
CD4 proliferating	1.82% (1.19%–2.67%)	2.8% (2.07%–3.75%)
Treg (CD4+)	2.63% (1.63%–4.14%)	1.37% (1.06%–1.67%)
CD4 resident memory	2.18% (1.02%–4.57%)	1.76% (0.968%–2.83%)

Table 4.21: Predicted composition by adjuvant with 95% credible intervals (KLH1 model 4a).

cell_group	KLH_montanide	KLH_alhydrogel
CD4 naive	91.4% (88.2%–93.3%)	92.9% (90.9%–94.5%)
CD4 PD-1+	1.92% (1.39%–2.68%)	1.23% (0.877%– 1.7%)
CD4 proliferating	1.83% (1.16%–2.92%)	1.39% (0.869%–2.17%)
Treg (CD4+)	2.61% (1.65%– 4.1%)	2.38% (1.37%– 3.7%)
CD4 resident memory	2.29% (1.03%–4.64%)	2.08% (0.947%–3.73%)

Table 4.22: Predicted composition by dose with 95% credible intervals (KLH1 model 4b).

cell_group	1	100
CD4\(^{+}\) TF\(^{-}\)	81.5% (79.1%–83.8%)	79.6% (77.7%–81.2%)
CD4\(^{+}\) GATA3\(^{+}\)	6.84% (5.36%–8.48%)	7.52% (6.18%–9.31%)
CD4\(^{+}\) GATA3\(^{+}\) TBET\(^{+}\)	0.926% (0.631%–1.34%)	1% (0.748%–1.32%)
CD4\(^{+}\) ROR\(\gamma\)t\(^{+}\)	1.61% (1.12%– 2.3%)	0.92% (0.795%–1.05%)
CD4\(^{+}\) TBET\(^{+}\) CXCR3\(^{+}\)	9.1% (7.27%–11.1%)	10.9% (9.47%–12.6%)

Table 4.23: Predicted composition by adjuvant with 95% credible intervals (KLH1 model 4b).

cell_group	KLH_montanide	KLH_alhydrogel
CD4\(^{+}\) TF\(^{-}\)	81.5% (79.2%–83.7%)	86.1% (83.8%–87.9%)
CD4\(^{+}\) GATA3\(^{+}\)	6.83% (5.34%–8.52%)	6.07% (4.69%–7.98%)
CD4\(^{+}\) GATA3\(^{+}\) TBET\(^{+}\)	0.909% (0.63%–1.34%)	0.487% (0.329%–0.745%)
CD4\(^{+}\) ROR\(\gamma\)t\(^{+}\)	1.61% (1.12%–2.36%)	1.4% (0.916%–2.08%)
CD4\(^{+}\) TBET\(^{+}\) CXCR3\(^{+}\)	9.14% (7.4%–11.1%)	5.99% (4.69%–7.65%)

4.6.5 Additional compositional models

Figure 4.32 examines biopsy composition according to primary immunisation regimen in control biopsies only. A statistically significant difference in the proportion of CD45 cells was observed in the placebo group; however, the magnitude of this effect was small. The absence of significant compositional differences for other immune cell populations suggests that control biopsies represent true baseline tissue and are not materially influenced by local or systemic effects of adjacent active rechallenge sites.

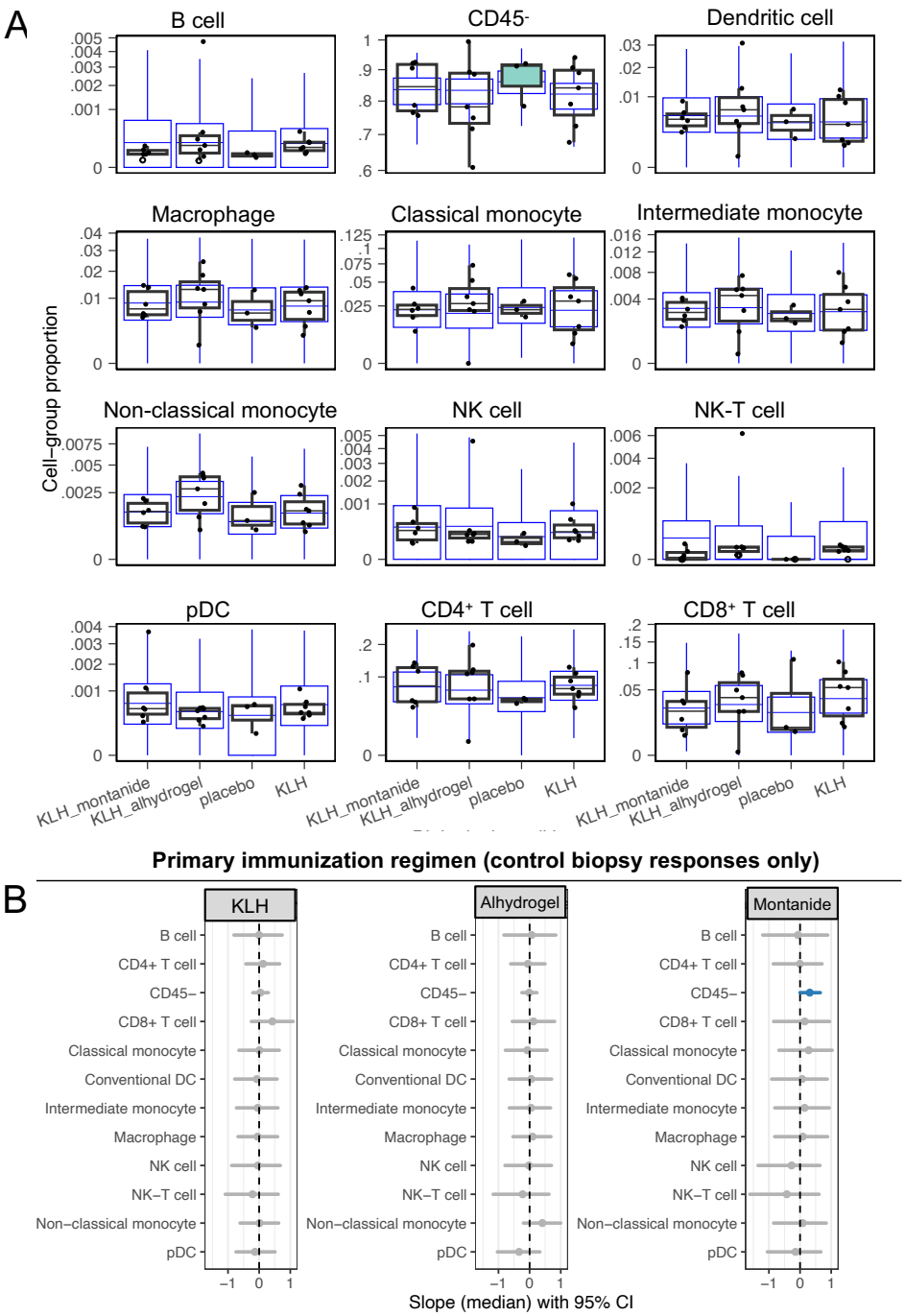

Figure 4.32: Compositional modelling results for **control** biopsy sites stratified by primary immunisation regimen in the KLH1 study. **Only control sites (0.9% saline-injected)are included**.
(A) Observed cell-type proportions (black boxplots) with posterior predictive distributions (blue boxplots) and individual data points, derived from Bayesian compositional models. Filled boxes indicate a statistically significant association between primary immunisation regimen and expected cell-type proportion at a 5% FDR. (B) Posterior mean estimates and 95% credible intervals for contrasts in expected cell-type proportion (logit scale) between unadjuvanted KLH, KLH + Alhydrogel, and KLH + Montanide relative to the placebo group. Intervals are coloured blue when statistically significant at a 5% FDR and grey otherwise.

4.6.6 Diagnostic information for compositional models

Review of the cmdstanr diagnostics suggested satisfactory tree depth for all transition, and no evidence of divergent transitions. E-BFMI was satisfactory, as were the effective sample size and R-hat values. Sampling adequacy and convergence for all compositional models were assessed using standard Hamiltonian Monte Carlo (HMC) diagnostics (Figures 4.33–4.38).

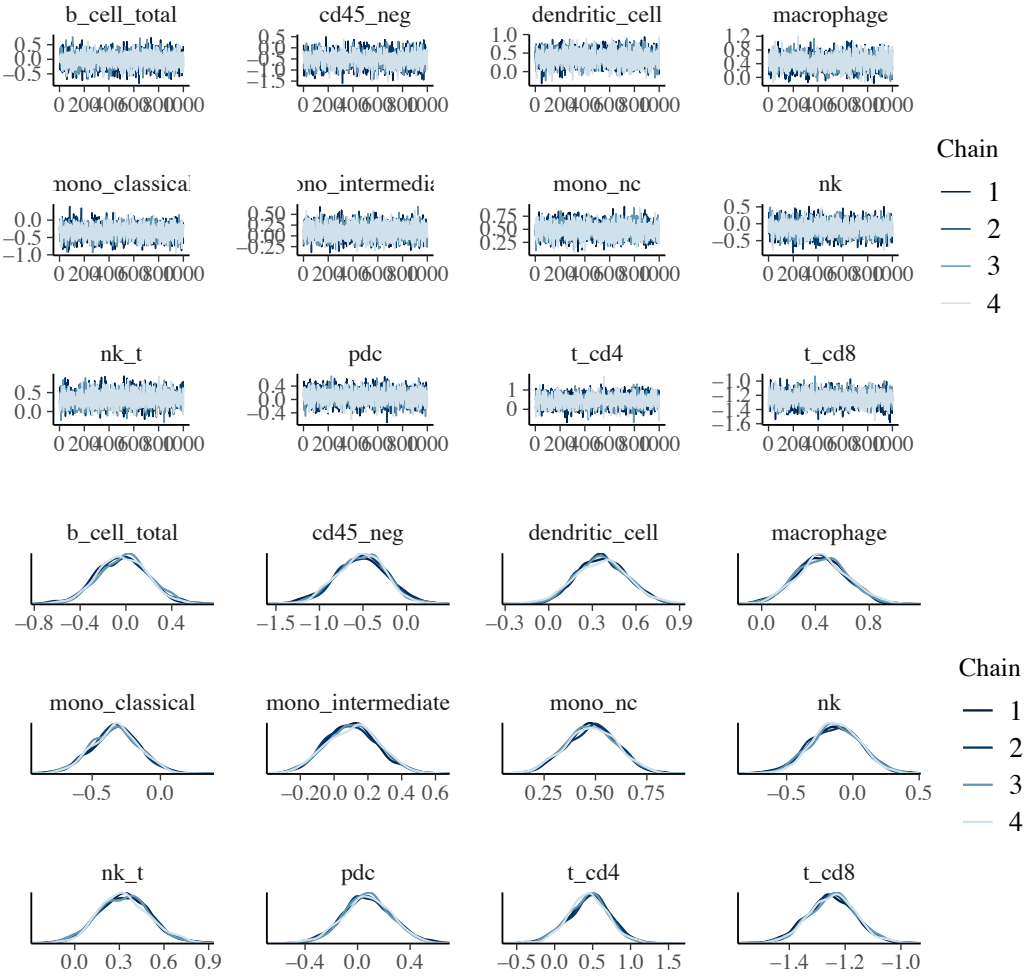

Figure 4.33: Compositional modelling diagnostics for Model 1 (induration status vs composition). Trace and marginal density plots indicate satisfactory Hamiltonian Monte Carlo (HMC) chain mixing and convergence.

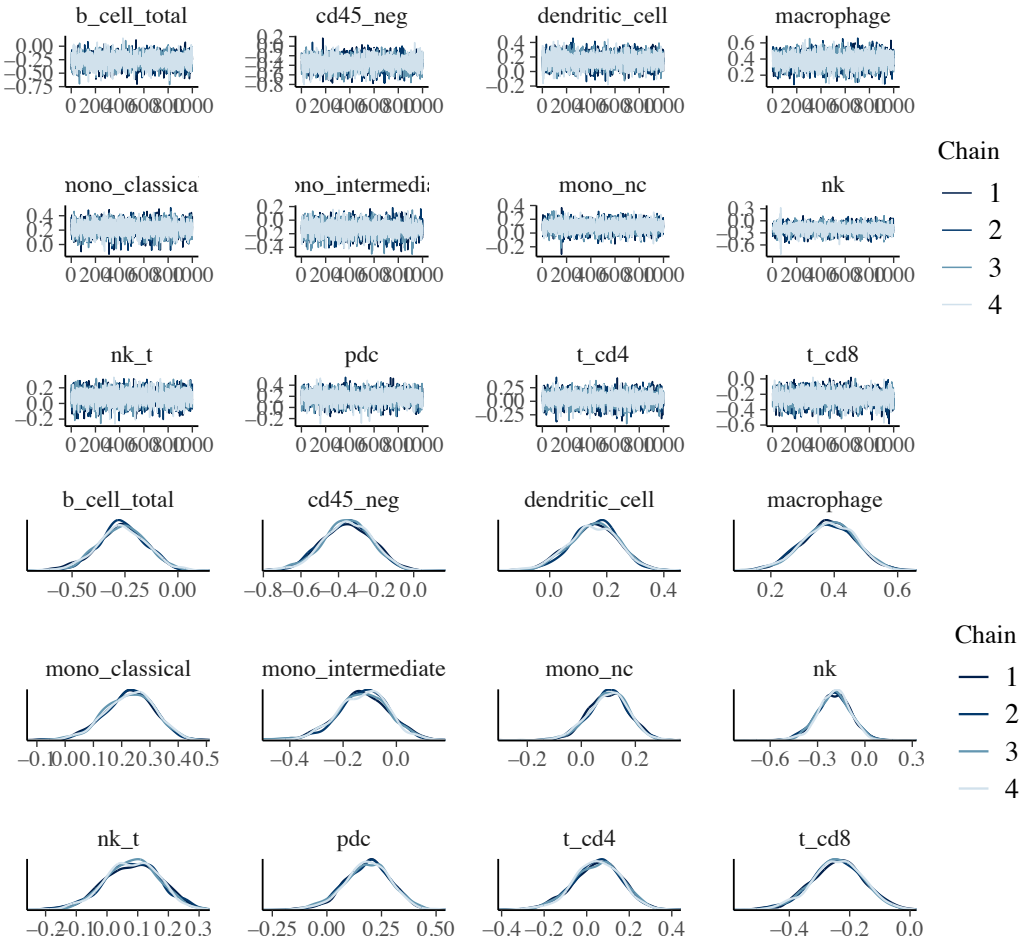

Figure 4.34: Compositional modelling diagnostics for Model 2 (induration diameter vs composition). Trace and marginal density plots indicate satisfactory Hamiltonian Monte Carlo (HMC) chain mixing and convergence.

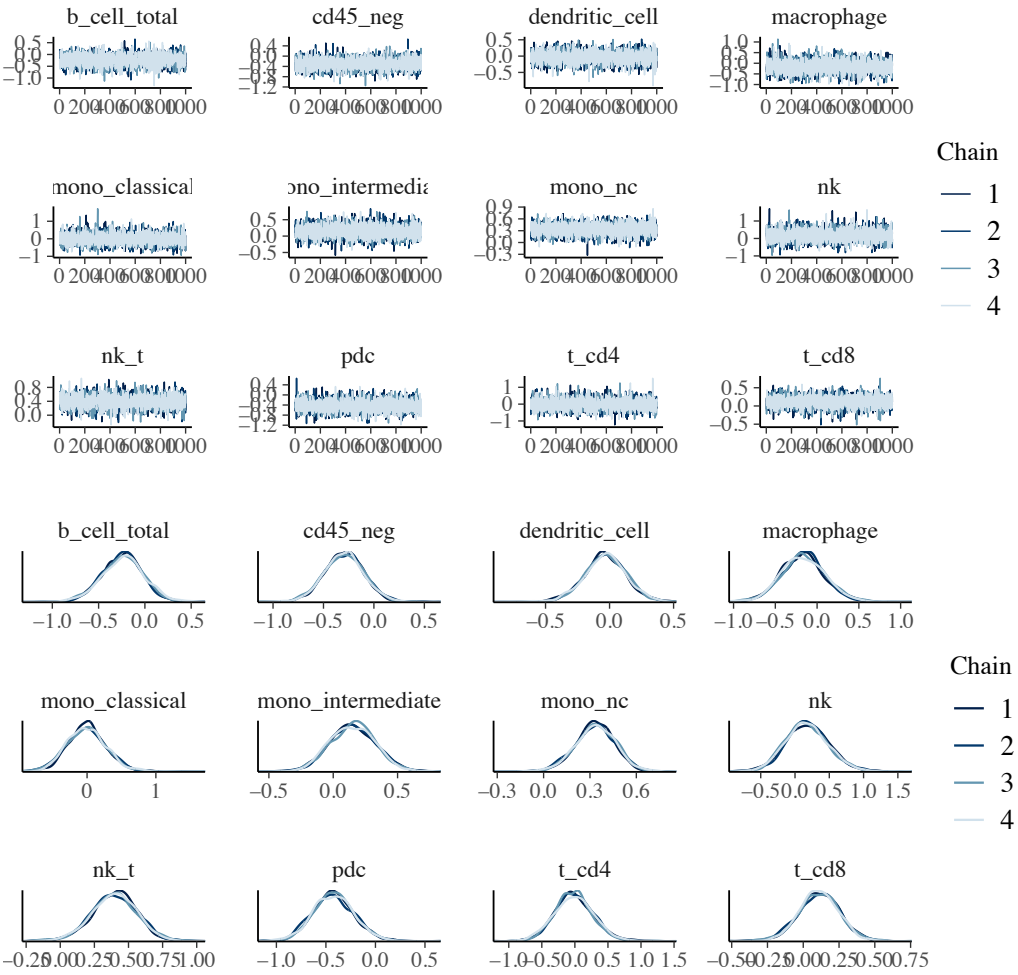

Figure 4.35: Compositional modelling diagnostics for Model 3 (dose and primary adjuvant vs cell-type composition). Trace and marginal density plots indicate satisfactory Hamiltonian Monte Carlo (HMC) chain mixing and convergence.

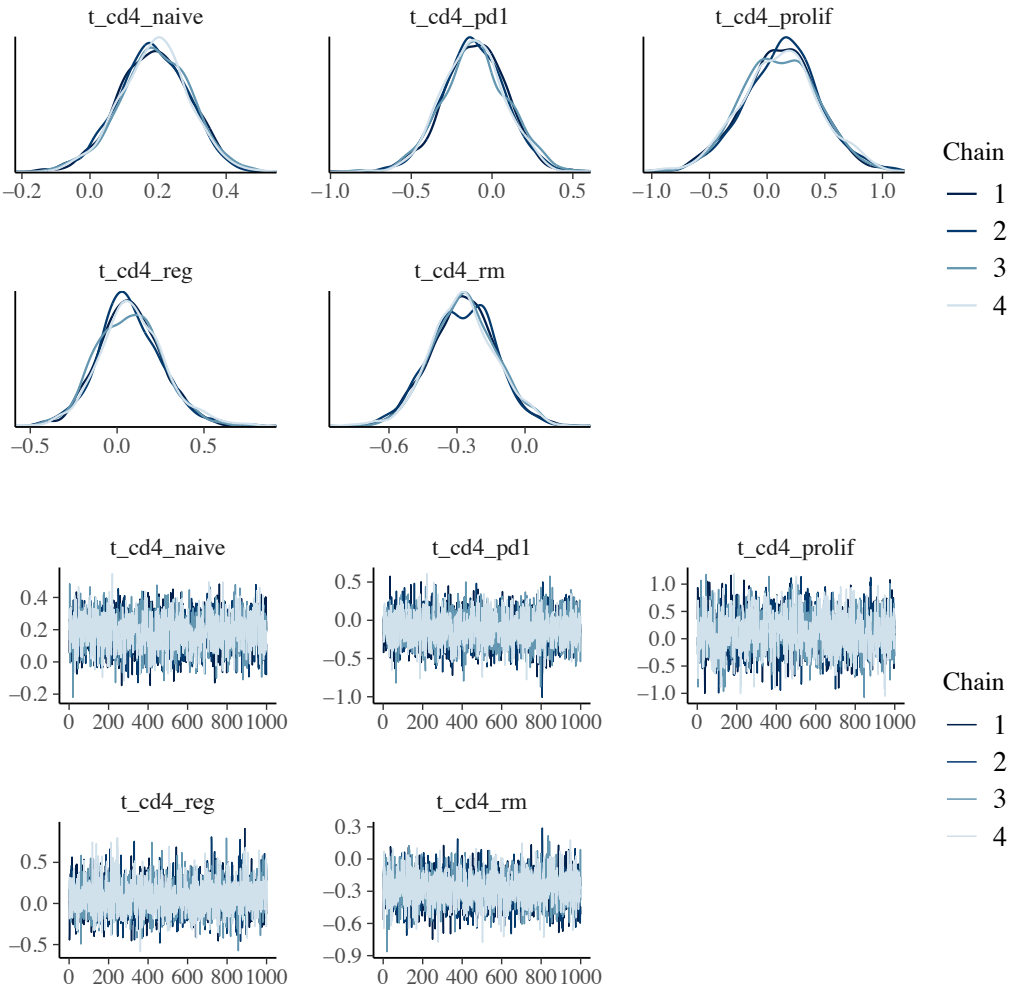

Figure 4.36: Compositional modelling diagnostics for Model 4a (dose and primary adjuvant vs CD4$^+$ T cell phenotypes). Trace and marginal density plots indicate satisfactory Hamiltonian Monte Carlo (HMC) chain mixing and convergence.

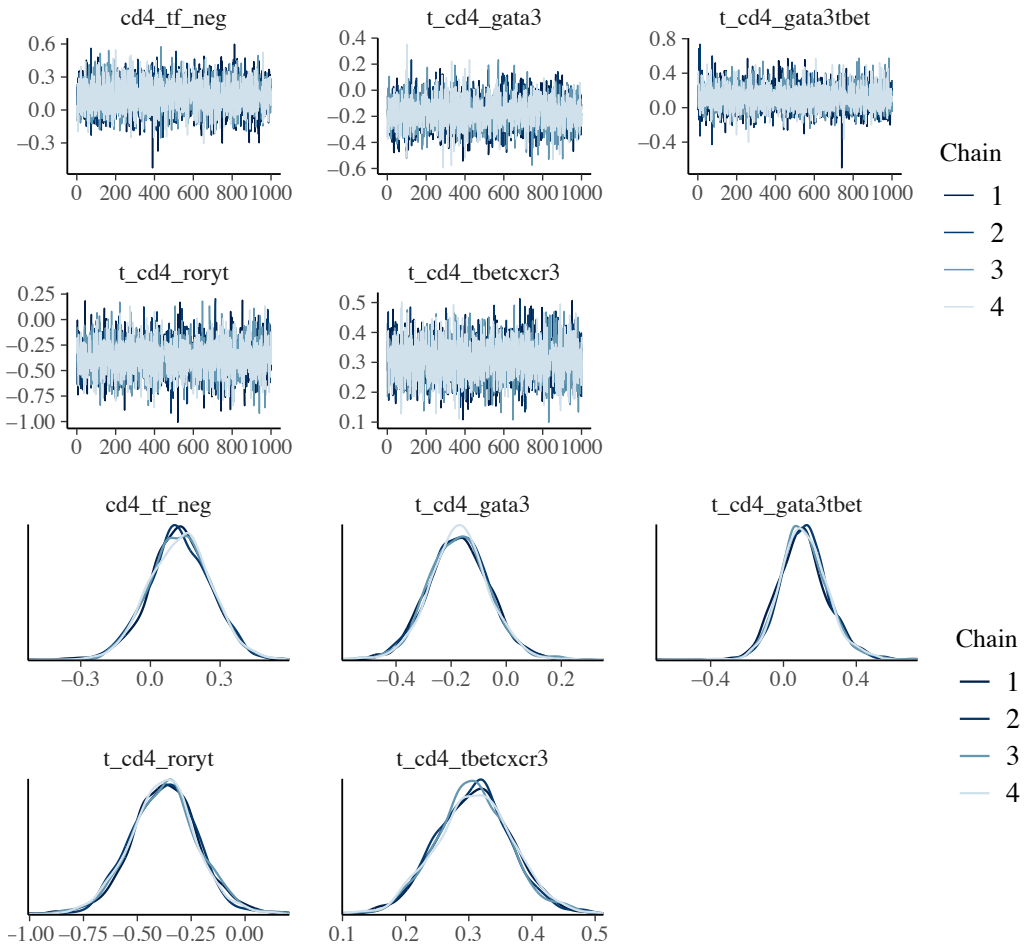

Figure 4.37: Compositional modelling diagnostics for Model 4b (dose and primary adjuvant vs CD4$^+$ T cell transcription factor expression). Trace and marginal density plots indicate satisfactory Hamiltonian Monte Carlo (HMC) chain mixing and convergence.

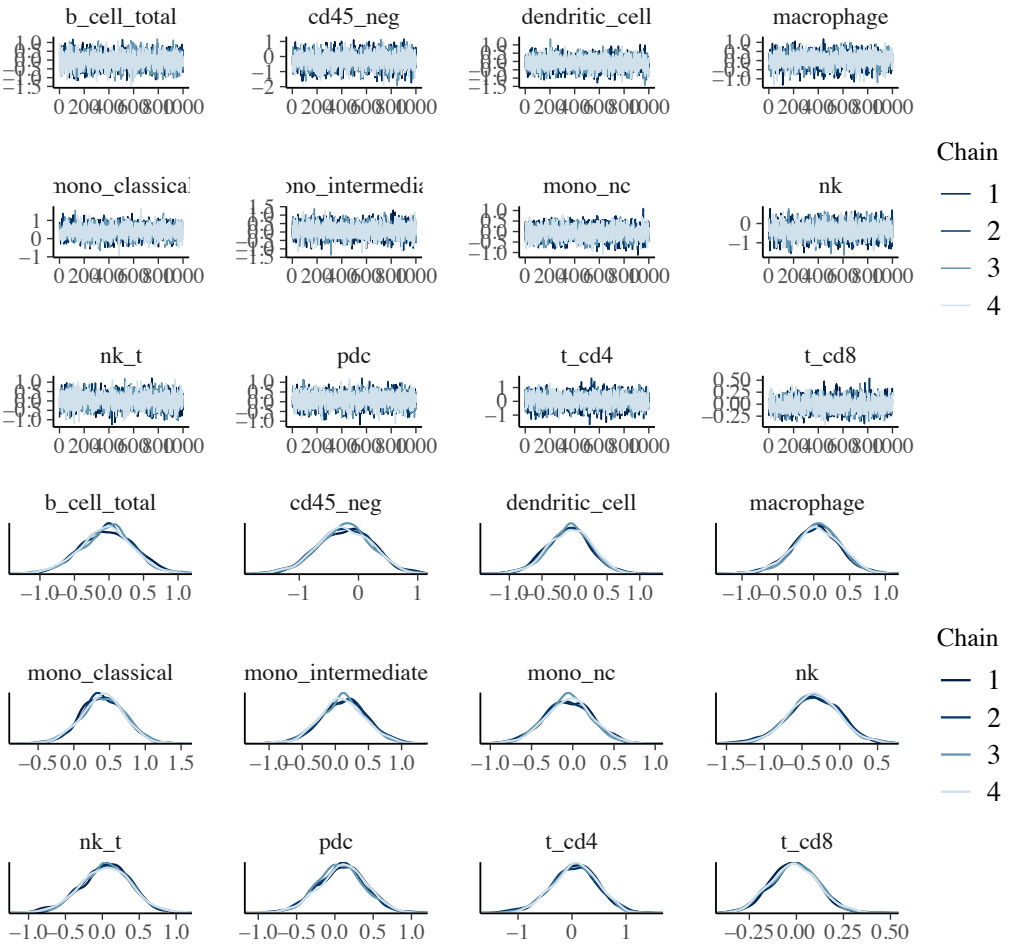

Figure 4.38: Compositional modelling diagnostics for Model 5 (primary immunisation regimen vs cell-type composition at control sites only). Trace and marginal density plots indicate satisfactory Hamiltonian Monte Carlo (HMC) chain mixing and convergence.

Table 4.24: Immune-related gene sets used for singscore analysis, grouped by source collection.

Gene set	Size (expressed genes)
Hallmark immune signatures (H)	
HALLMARK_ALLOGRAFT_REJECTION	182
HALLMARK_APOPTOSIS	153
HALLMARK_COMPLEMENT	179
HALLMARK_IL2_STAT5_SIGNALING	185
HALLMARK_IL6_JAK_STAT3_SIGNALING	82
HALLMARK_INFLAMMATORY_RESPONSE	181
HALLMARK_INTERFERON_ALPHA_RESPONSE	95
HALLMARK_INTERFERON_GAMMA_RESPONSE	194
HALLMARK_PI3K_AKT_MTOR_SIGNALING	97
HALLMARK_TNFA_SIGNALING_VIA_NFKB	197
KEGG immune pathways (C2:CP:KEGG)	
KEGG_B_CELL_RECEPTOR_SIGNALING_PATHWAY	73
KEGG_CHEMOKINE_SIGNALING_PATHWAY	156
KEGG_COMPLEMENT_AND_COAGULATION_CASCADES	43
KEGG_CYTOKINE_CYTOKINE_RECEPTOR_INTERACTION	191
KEGG_JAK_STAT_SIGNALING_PATHWAY	112
KEGG_NATURAL_KILLER_CELL_MEDIATED_CYTOTOXICITY	108
KEGG_NOD_LIKE_RECEPTOR_SIGNALING_PATHWAY	57
KEGG_RIG_I_LIKE_RECEPTOR_SIGNALING_PATHWAY	48
KEGG_TOLL_LIKE_RECEPTOR_SIGNALING_PATHWAY	84
KEGG_T_CELL_RECEPTOR_SIGNALING_PATHWAY	103
Reactome immune pathways (C2:CP:REACTOME)	
REACTOME_ADORA2B_MEDIATED_ANTI_INFLAMMATORY_CYTOKINES_PRODUCTION	32
REACTOME_ANTIGEN_ACTIVATES_B_CELL_RECEPTOR_BCR_LEADING_TO_GENERATION_OF_SECOND_MESSENGERS	30
REACTOME_ANTIGEN_PRESENTATION_FOLDING_ASSEMBLY_AND_PEPTIDE_LOADING_OF_CLASS_I_MHC	29
REACTOME_ANTIGEN_PROCESSING_CROSS_PRESENTATION	92
REACTOME_ANTIGEN_PROCESSING_UBIQUITINATION_PROTEASOME_DEGRADATION	263
REACTOME_ANTIGEN_PROCESSING_UB_ATP_INDEPENDENT_PROTEASOMAL_DEGRADATION	19
REACTOME_CHEMOKINE_RECEPTORS_BIND_CHEMOKINES	46
REACTOME_CLASS_I_MHC_MEDIATED_ANTIGEN_PROCESSING_PRESENTATION	336
REACTOME_COMPLEMENT_CASCADE	32
REACTOME_CROSS_PRESENTATION_OF_SOLUBLE_EXOGENOUS_ANTIGENS_ENDOSOMES	40
REACTOME_DDX58_IFIH1_MEDIATED_INDUCTION_OF_INTERFERON_ALPHA_BETA	64
REACTOME_DISEASES_OF_IMMUNE_SYSTEM	28
REACTOME_DOWNSTREAM_SIGNALING_EVENTS_OF_B_CELL_RECEPTOR_BCR	67
REACTOME_DOWNSTREAM_TCR_SIGNALING	79
REACTOME_EVASION_BY_RSV_OF_HOST_INTERFERON_RESPONSES	18
REACTOME_IL_6_TYPE_CYTOKINE_RECEPTOR_LIGAND_INTERACTIONS	11
REACTOME_INITIAL_TRIGGERING_OF_COMPLEMENT	16
REACTOME_INTERFERON_ALPHA_BETA_SIGNALING	64
REACTOME_INTERFERON_GAMMA_SIGNALING	86
REACTOME_INTERFERON_SIGNALING	236
REACTOME_MHC_CLASS_II_ANTIGEN_PRESENTATION	107
REACTOME_MITOTIC_TELOPHASE_CYTOKINESIS	12
REACTOME_NEF_MEDIATED_DOWNREGULATION_OF_MHC_CLASS_I_COMPLEX_CELL_SURFACE_EXPRESSION	10
REACTOME_NEUTROPHIL_DEGRANULATION	416
REACTOME_REGULATION_OF_INNATE_IMMUNE_RESPONSES_TO_CYTOSOLIC_DNA	13
REACTOME_SARS_COV_1_ACTIVATES_MODULATES_INNATE_IMMUNE_RESPONSES	36
REACTOME_SARS_COV_2_ACTIVATES_MODULATES_INNATE_AND_ADAPTIVE_IMMUNE_RESPONSES	100
REACTOME_SIGNALING_BY_THE_B_CELL_RECEPTOR_BCR	96
REACTOME_STING_MEDIATED_INDUCTION_OF_HOST_IMMUNE_RESPONSES	12
REACTOME_SUMOYLATION_OF_IMMUNE_RESPONSE_PROTEINS	12
REACTOME_TCR_SIGNALING	100
REACTOME_TOLL_LIKE_RECEPTOR_9_TLR9_CASCADE	103
REACTOME_TOLL_LIKE_RECEPTOR_CASCADES	154

Table 4.24: Immune-related gene sets used for singscore analysis, grouped by source collection. *(continued)*

Gene set	Size (expressed genes)
REACTOME_TOLL_LIKE_RECEPTOR_TLR1_TLR2_CASCADE	105
Transcription factor targets (C3:TFT_LEGACY)	
GATA3_01	167
NFAT_Q4_01	207
NFAT_Q6	185

Chapter 5

Optimising the identification of primary draining lymph nodes using contrast enhanced ultrasound: a pilot study

Chapter summary

Building on the need for tissue-level pharmacodynamic readouts in human immune challenge and vaccine studies, this chapter evaluates the feasibility of CEUS to improve identification of primary draining lymph nodes (LNs) for fine needle aspiration. Accurate localisation of the draining LN is critical for reliable interpretation of LN-based immune readouts, yet remains technically challenging using conventional ultrasound alone.

This chapter reports a pilot experimental medicine study assessing subcutaneous and intramuscular administration of Sonazoid microbubble contrast at common antigen delivery sites, using an adaptive Bayesian monitoring framework to guide early stopping decisions.

This work addresses Objective 4 of the thesis:

To evaluate, using an experimental medicine approach, the feasibility of contrast-enhanced ultrasound for identifying primary draining lymph nodes in healthy volunteers to enable tissue-based sampling in HIC and PoM studies.

Despite clear contrast enhancement at injection sites and occasional visualisation of lymphatic channels, CEUS did not reliably identify primary draining LNs amenable to aspiration, lead-

ing to early termination of the study for futility.

5.1 Introduction

The lymph node (LN) is the principal coordinating site for adaptive immune responses to exogenous antigens and is therefore of central interest for both fundamental immunology and translational studies assessing immunomodulatory agents.[114] Advances in single-cell and molecular techniques now allow exploration of LN biology at unprecedented resolution.

LN sampling, typically via ultrasound-guided FNA, is safe and well-tolerated. Recent studies have demonstrated the feasibility of LN FNA in healthy volunteers for vaccine and immune challenge studies.[115, 116] However, the wider utility of LN FNA remains constrained by difficulties in reliably identifying the true primary draining LN for sampling. Inadequate targeting can lead to false-negative readouts of antigenic activation, undermining the interpretability of immunological endpoints. For example, in a study of SARS-CoV-2 vaccination, activated LN cells were detected in only 5/8 participants, likely reflecting failure to sample the true draining LN.[117] PET-CT imaging studies further highlight the high inter-individual variability of lymphatic drainage from common injection sites and the frequent involvement of deep LNs not amenable to FNA.[118]

5.1.1 Microbubble CEUS for identification of draining lymph nodes

Microbubble contrast agents consist of small (2 µm to 10 µm) particles of inert gas surrounded by an encapsulating lipid shell. Following administration, microbubbles oscillate in response to interaction with ultrasound waves, leading to enhancement relative to surrounding tissues.[119, 120] Microbubble contrast agents delivered by intravenous injection are licensed and widely used to improve visibility of cardiac chambers and characterisation of liver lesions. Microbubbles administered intradermally have been shown to transit to lymph nodes via lymphatic channels directly, and following phagocytosis by tissue macrophages,[121] leading to their use for mapping of lymphatics and identification of sentinel lymph nodes in breast cancer and other tumour types.[122–124] It has been

noted that some microbubble contrast products, including Sonazoid (GE Healthcare, Oslo, Norway), are efficiently phagocytosed by cells of the reticuloendothelial system, and are thus well-suited to the imaging of secondary lymphoid organs such as lymph nodes. Following percutaneous administration of Sonazoid, enhancement of lymph nodes occurs rapidly (within minutes), and has been shown to persist for up to 24 hours.[125]

Although microbubble CEUS has demonstrated clinical utility for identifying lymph nodes in patient groups,[121, 122] and animals,[123] it has not to our knowledge previously been used as a tool in human experimental medicine. The established safety and tolerability profile of the approach render it well-suited to healthy volunteer studies. This study will provide further data to characterise the potential role of this technique in a research setting.

The aim of the study was therefore to determine whether CEUS utilizing Sonazoid microbubbles administered to different sites (deltoid, forearm, and lateral thigh) and routes (subcutaneous and intramuscular) commonly used for immune challenge administration enables reliable identification of primary draining lymph nodes from these areas. Specific objectives of this study were

Chapter Objective(s)

1. To identify and map primary draining lymph nodes in healthy volunteers using contrast-enhanced ultrasound (CEUS, Sonazoid, GE Healthcare) following subcutaneous or intramuscular administration, at different anatomical sites,
2. To perform in-depth assessment of performance characteristics of CEUS for identification of primary draining lymph nodes,
3. To determine the feasibility of CEUS for elucidation of basic physiology and immunology in healthy volunteers

5.2 Methods

5.2.1 Study design

This was a non-blinded, single-arm experimental medicine study designed to evaluate CEUS using the microbubble contrast agent Sonazoid for identification of primary draining LNs in healthy adult volunteers. The study followed an adaptive Bayesian design implemented using the `Bayesian Efficacy Monitoring via Predictive Probability` (BEMPR) framework (v1.1.4.0); `https://trialdesign.org/one-page-shell.html#BEMPR`).

An overview of the study design and procedures is shown in Figure 5.1.

5.2.2 Participants

Healthy male or female adults aged 18–45 years with a body mass $\geq$50kg and BMI between 18–30kg/m^2 were eligible. Exclusion criteria included medical conditions affecting immune or lymphatic function, recent immunosuppressive therapy, or contraindications to Sonazoid (e.g. egg allergy). Full eligibility criteria were detailed in the study protocol. A total of four participants were enrolled prior to early termination.

5.2.3 Intervention and procedures

Each participant underwent up to five contrast-enhanced ultrasound (CEUS) procedures over two study visits conducted at least seven days apart (Figures 5.1 and 5.2). Sonazoid was reconstituted and administered according to the manufacturer's instructions using a 21-gauge needle. CEUS examinations were performed on a GE Logiq E10 ultrasound system using a high-frequency linear array transducer (4–20 MHz) and a standardised imaging protocol incorporating vendor-recommended settings for contrast-specific imaging. A mechanical index (MI) of 0.14 was used for continuous imaging, with brief increases permitted where necessary to optimise contrast visualisation.

For each procedure, baseline (non-contrast) ultrasound assessment was performed to document lymph node morphology and the presence of visible lymphatic channels. Following Sonazoid ad-

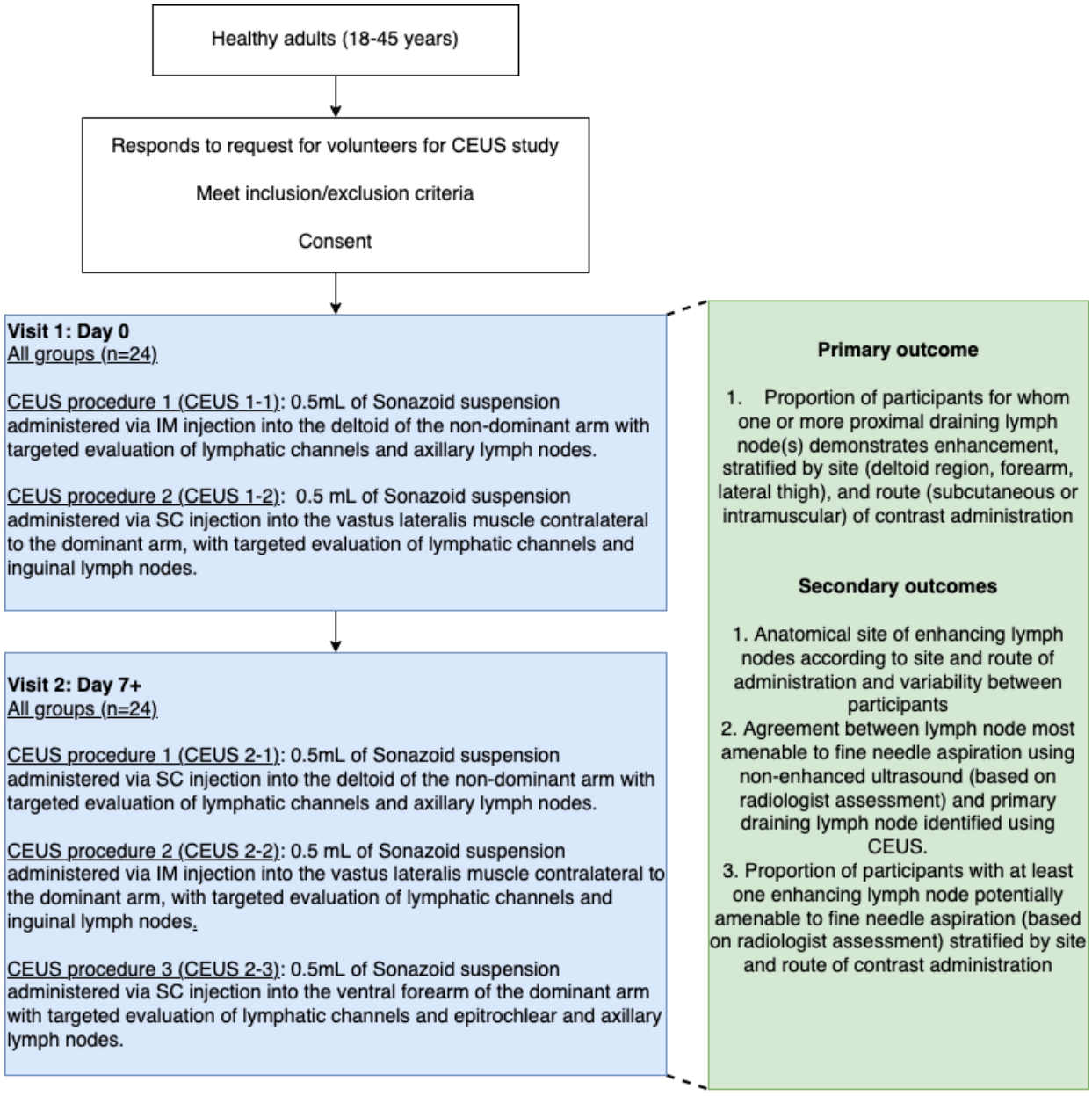

Figure 5.1: Overview of CEUS study design and key objectives.

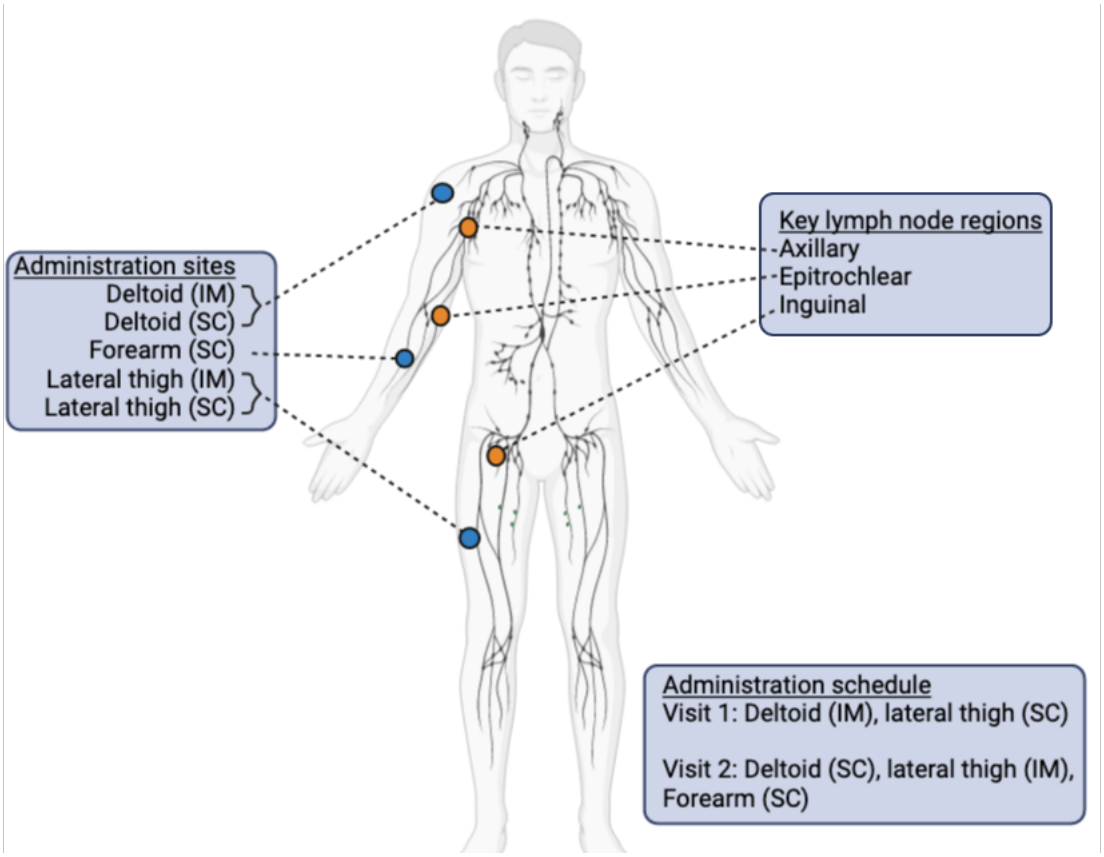

Figure 5.2: Sites of Sonazoid administration and corresponding lymph node groups for ultrasound assessment.

ministration, the injection site was massaged thoroughly for a minimum of one minute to facilitate dispersion and lymphatic uptake of contrast. If, in the opinion of the scanning clinician, contrast enhancement of lymphatic channels or lymph nodes was insufficient, a second dose of Sonazoid (0.5 mL) was administered at the same site, with reconstitution performed according to the manufacturer's instructions.

Data collected for each CEUS procedure included visibility of lymphatic channels, time to contrast enhancement, and qualitative assessment of enhancement. Still images and cine loop recordings were archived for subsequent review.

5.2.4 Adaptive Bayesian design and data monitoring

The study employed an adaptive Bayesian monitoring framework (BEMPR v1.1.4.0) to allow early stopping for futility or efficacy based on predictive probability. Input parameters were as follows:

This framework allowed the study to adaptively terminate if the posterior probability of achieving the efficacy threshold fell below 0.1 (futility) or exceeded 0.95 (efficacy). The primary endpoint—successful identification of a contrast-enhancing primary draining LN amenable to FNA—was modelled as a binary response.

5.2.5 Statistical methods

All analyses were descriptive. Categorical variables were summarised as proportions (with 95% binomial confidence intervals), and continuous variables as means (95% CI) or medians (interquartile range), as appropriate. Logistic regression was planned to explore associations between injection route, site, and LN identification success, contingent on sufficient sample size.

A

Bayesian Efficacy Monitoring Via Predictive Probability v1.1.3.0	
Stopping Boundaries Input	
Trial Name (Optional)	Microbubbles
Reference response rate (theta0)	0.4
Threshold for declaring efficacy at the end of the trial (pT)	0.95
Early Stopping for Futility	TRUE
Probability confidence threshold for futility stopping (pL)	0.1
Early Stopping for Efficacy	TRUE
Probability confidence threshold for declaring efficacy early (pU)	0.95
Prior distribution for theta:Beta(a0,b0) a0	1
Prior distribution for theta:Beta(a0,b0) b0	1
Input Cohort Manually	FALSE
Maximum number of patients in the trial	24
Minimum number of patients before early stopping rule applies	4
Cohort size	4

B

# Patients (inclusive)	# Responses (inclusive) are considered efficacious	Actions
4	4	Early stopping
8	7 – 8	Early stopping
12	9 – 12	Early stopping
16	11 – 16	Early stopping
20	13 – 20	Early stopping

C

# Patients (inclusive)	# Responses (inclusive) are considered futile	Actions
4	0 – 1	Early stopping
8	0 – 3	Early stopping
12	0 – 5	Early stopping
16	0 – 7	Early stopping
20	0 – 10	Early stopping

Figure 5.3: Bayesian monitoring framework for dynamic sample size estimation based on predictive probability of success. A: Targets and assumptions for monitoring plan (inputs to BEMPR v1.1.4.0). B: Outcome scenarios leading to early stopping for efficacy, based on interim assessment of cohorts sized n=4. C: Outcome scenarios leading to early stopping for futility. Efficacy was defined as identification of at least one contrast-enhancing primary draining lymph node at a corresponding anatomical location following Sonazoid administration at any of the administration sites

5.3 Results

5.3.1 Participant characteristics

Four participants (3 male, 1 female; median age 29 years) were enrolled prior to early termination of the study.

5.3.2 Adaptive monitoring and study termination

At the first interim analysis ($n = 4$), no participants demonstrated clear CEUS enhancement of a draining LN amenable to FNA from the target injection sites at least 30 minutes post contrast injection. Based on the Bayesian predictive probability framework, the posterior probability of ultimate success fell below the futility threshold. The study was therefore terminated early for futility in accordance with the adaptive design.

5.3.3 CEUS performance and kinetics

In all participants, Sonazoid administration was well tolerated with no adverse events. Lymph nodes could be identified prior to contrast injection at axillary and inguinal sites anatomically adjacent to the target injection site in all four participants (figure 5.4). Administration of Sonazoid resulted in immediate, high-intensity contrast enhancement at the site of injection, confirming the echogenicity of the reconstituted microbubbles (figure 5.5). Enhancement persisted for the duration of the imaging procedure with minimal attenuation over time.

In one participant, a discrete structure potentially consistent with a lymphatic channel was visualised (Figure 5.6). This appeared as a short linear, hyperechoic structure extending proximally from the injection depot with subtle undulation of signal consistent with microbubble transit.

In the remaining three participants, contrast enhancement remained strictly confined to the immediate injection depot. No egress of microbubbles into draining channels was observed despite observation for a minimum of 30 minutes in conjunction with active limb flexion and massage of the site. Consequently, no candidate primary draining lymph nodes were identified in any participant.

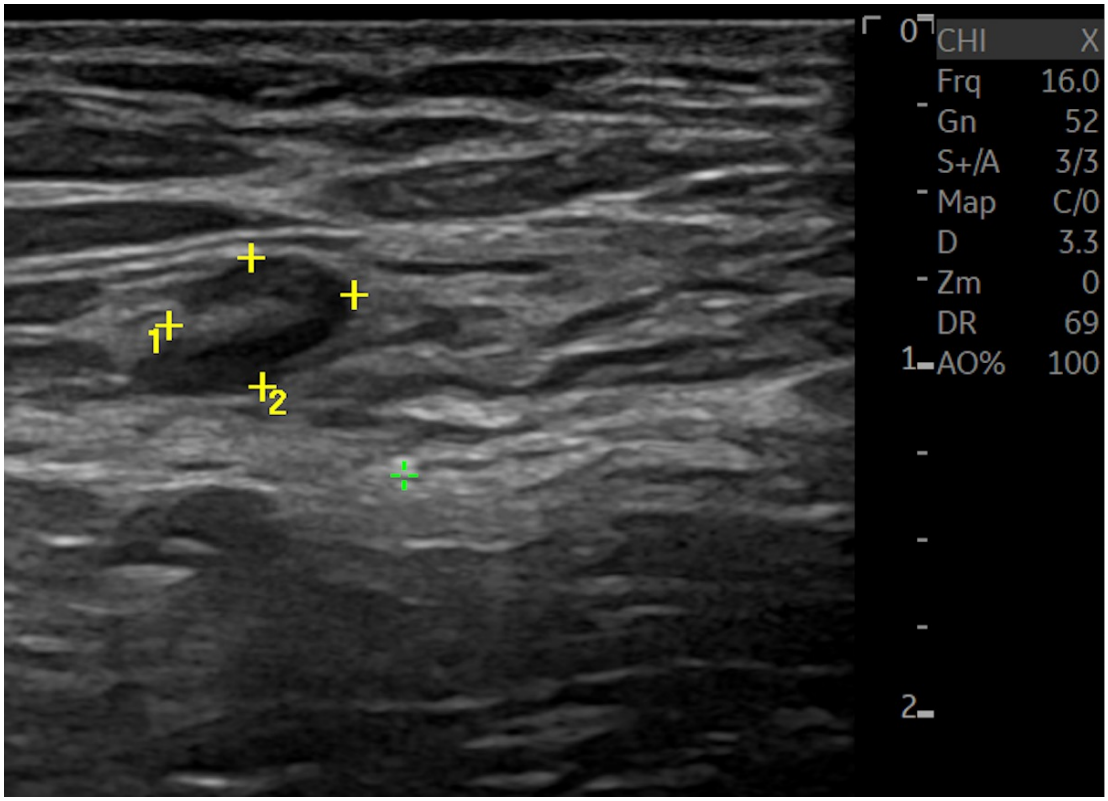

Figure 5.4: Example of a non-enhancing axillary lymph node. Imaging performed by Dr Paul Lyon.

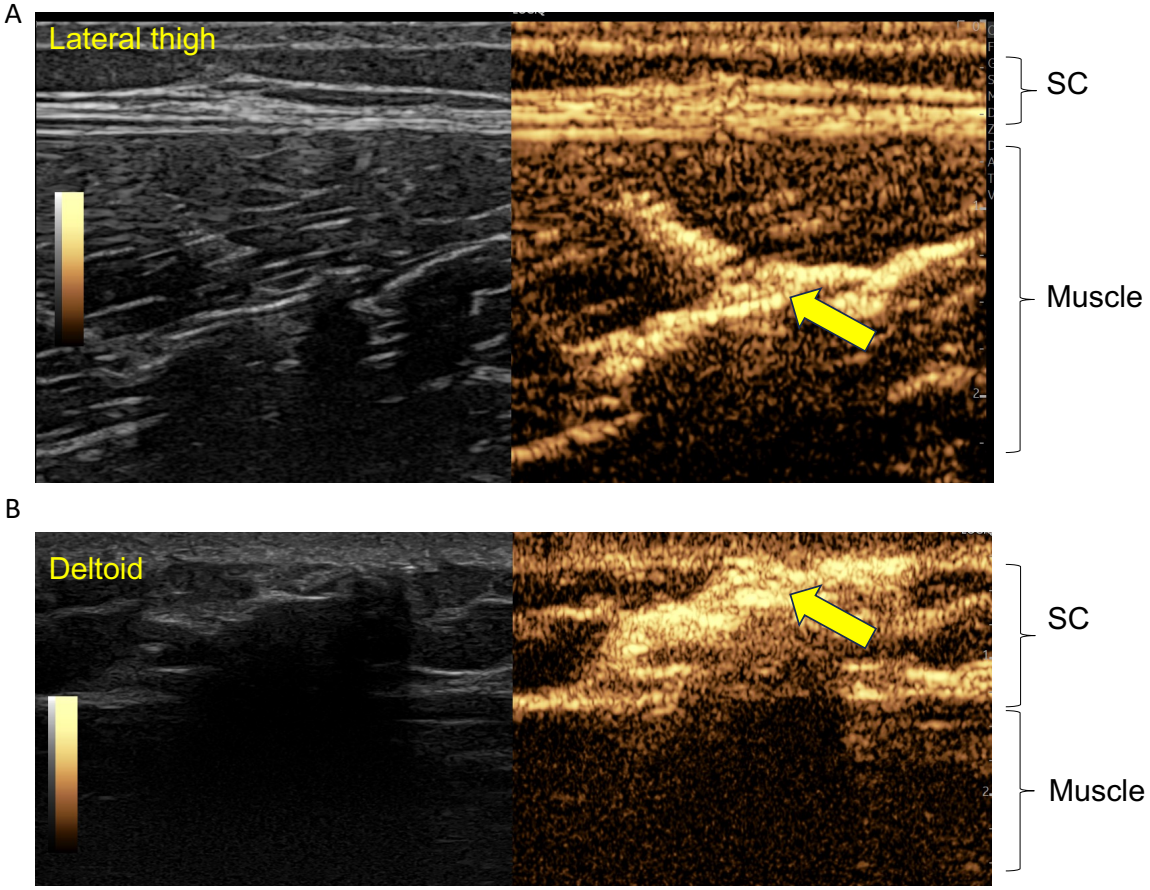

Figure 5.5: Exemplary CEUS images showing contrast appearance at the site of injection: (A) lateral thigh, intramuscular; (B) deltoid, subcutaneous. Sites of contrast deposition are indicated by yellow arrows. Imaging performed by Dr Paul Lyon.

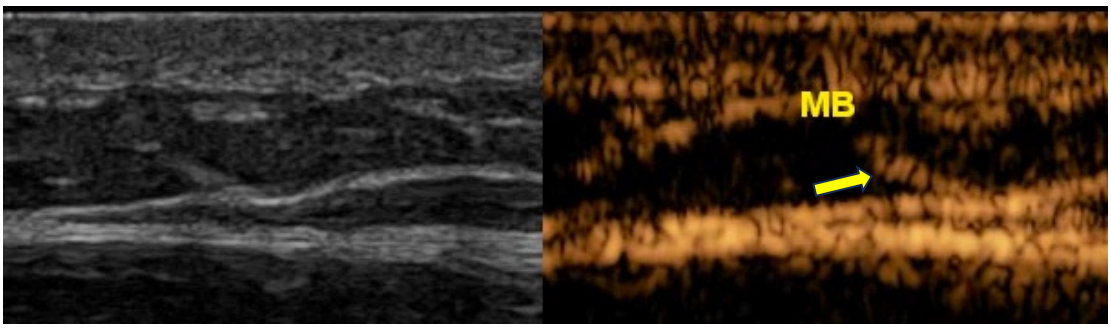

Figure 5.6: CEUS image demonstrating a discrete enhancing lymphatic channel in the right deltoid region following subcutaneous Sonazoid administration. The channel is indicated by a yellow arrow. Imaging performed by Dr Paul Lyon.

5.4 Discussion

5.4.1 Key findings

This experimental medicine study evaluated the feasibility of CEUS for identifying primary draining lymph nodes (LNs) following immune challenge site injections in healthy volunteers. Using an adaptive Bayesian monitoring framework, the study was terminated early after enrolment of four participants due to futility, indicating that CEUS, as implemented here, did not reliably identify primary draining LNs amenable to fine needle aspiration.

The absence of identifiable contrast-enhancing lymph nodes contrasts with multiple reports of successful sentinel LN visualisation following Sonazoid administration in breast tissue and oncological settings.[126–128] Several factors may account for this discrepancy. First, the anatomical characteristics of the administration sites used in this study differ substantially from breast tissue: the deltoid, forearm, and lateral thigh are characterised by relatively dense muscle and subcutaneous compartments, which may have impaired microbubble transit through lymphatic channels or reduced microbubble survival prior to LN uptake. Second, lymphatic flow from unstimulated immune challenge sites in healthy volunteers may be insufficient to generate detectable LN enhancement, particularly where baseline lymph node size is small. Third, it is possible that longer observation intervals following contrast administration would be required to permit lymphatic transport and nodal accumulation of microbubbles in this context. Finally, although all scans were performed by an experienced consultant radiologist, the absence of subspecialist expertise in CEUS-guided lymphatic imaging may have contributed to reduced technical sensitivity.

5.4.2 Strengths and limitations

The principal strength of this study was the use of an adaptive Bayesian design, which enabled data-driven early stopping in the setting of uncertain technical feasibility. This approach ensured efficient use of resources and minimised participant exposure while maintaining scientific rigour, and is particularly well suited to exploratory experimental medicine studies.

The primary limitation was the lack of opportunity for iterative protocol refinement. Time and resource constraints, together with the need for formal ethical approval for protocol amendments, precluded systematic exploration of alternative imaging strategies. It is therefore possible that modifications such as longer post-injection observation intervals (e.g. 24 h), administration in the context of an active immune challenge associated with lymph node reactivity, or alternative injection techniques may have improved lymph node visualisation.

An additional consideration is the potential for a learning effect associated with technically demanding lymphatic imaging procedures, particularly in anatomically complex regions such as the axilla. Although all scans were performed by an experienced consultant radiologist, it remains possible that technical success might have improved with greater procedural volume or subspecialist CEUS experience. However, given the absence of consistent lymph node enhancement across participants and injection sites, these findings raise questions regarding the suitability of CEUS as a general-purpose tool for lymph node identification in human experimental medicine settings without substantial further optimisation.

5.5 Conclusions

The findings suggest that, in the absence of substantial further technical optimisation, CEUS with Sonazoid has limited utility for identifying primary draining lymph nodes suitable for fine needle aspiration in healthy volunteers receiving subcutaneous or intramuscular immune challenges. While CEUS reliably produced robust contrast enhancement at injection sites, this did not translate into consistent visualisation of draining lymph nodes amenable to sampling.

Importantly, this study demonstrates the feasibility and efficiency of Bayesian adaptive methods for evaluating technical and methodological questions in early-phase experimental medicine. By enabling rapid, evidence-based termination of a low-yield approach, the adaptive design conserved resources and minimised participant exposure, directly supporting the broader thesis objective of optimising experimental paradigms to maximise decision value. In this context, the negative findings reported here are informative.

Chapter 6

A decision-theoretic approach to early-phase Proof-of-Mechanism studies

Chapter summary

In this chapter we move from considering the development and optimisation of immune challenge models, to their application in a Go/No-go decision context. Using a Bayesian decision-theoretic approach, we illustrate how PD response data from an HIC study can be formally integrated with utility metrics.

The work addresses objective 5 of the thesis:

To demonstrate how immune challenge-derived PD readouts may be formally integrated within a Bayesian decision framework, and to examine how utility-calibrated decision rules can optimise Go/No-Go thresholds relative to conventional Phase I decision strategies.

While intentionally simplified, the simulation study presented in this chapter demonstrates that integration of PD readouts into a quantitative decision-making framework is feasible. By transparently encoding and propagating uncertainty, this approach enables utility-optimal sample size estimation and probabilistic decision thresholds which may differ substantially from those selected based on qualitative heuristics or historical convention.

6.1 Introduction

Early-phase clinical development is characterised by decision-making under profound uncertainty in an environment of high economic stakes. Following a Phase I trial, decision makers face the task of evaluating whether the accumulated evidence justifies the substantial investment required to progress a compound to Phase II ("go"), or whether development should be curtailed or redirected ("no-go"). This decision balances both scientific plausibility (Is the drug working mechanistically as intended and likely to translate into clinical efficacy?) and economic viability (Is the expected return on investment sufficient?).

Much recent attention has been given to the importance of establishing PoM to inform Go/No-go decisions. One prominent articulation of a decision-making rubric for PoM assessment is Pfizer's 3-pillars paradigm, defined by evidence for exposure (of the drug) at the site of action (Pillar 1), binding to the pharmacological target (Pillar 2), and expression of pharmacological activity (Pillar 3).[129] In a 10 year retrospective analysis of the Pfizer pipeline, establishing PoM via the 3-pillars paradigm was credited with an improvement in the three year rolling average phase II-phase III transition success from 19% to 53%.[10] In this analysis, the association with a positive 3-pillar rating was stark, with 29% phase II success rates observed in these compounds between 2010 and 2019, compared to 0% with a negative 3-pillar rating.

However while straightforward in application (and communication), a categorical heuristic rule obscures substantial complexity, especially in the assessment of downstream pharmacological activity. Readouts of immune system processes, such as those elicited by HIC, are subject to substantial heterogeneity, mostly due to non-heritable factors.[29, 130] In the context of immunomodulatory drug development this unexplained heterogeneity constitutes statistical 'noise' which must be distinguished from the 'signal' of the drug effect being tested, and while model optimisation can address this issue to some extent, comprehensive optimisation prior to application is unlikely to be feasible in many development scenarios.

Furthermore, standard statistical approaches treat these biomarkers merely as numerical endpoints to be tested, rather than mechanistic biological signals which carry probabilistic (rather than

deterministic) information about downstream clinical success. A potential solution to these challenges is to transparently acknowledge and quantify uncertainty in the evidence base contributing to a Go/No-go decision through formal modelling, using the long-run economic consequences of the approach to calibrate decision thresholds. Bayesian Decision Theory provides a probabilistic framework for making optimal choices under uncertainty.[131] Unlike classical hypothesis testing, which focuses on controlling error rates, BDT integrates three core components: quantified uncertainty via the Bayesian posterior distribution, the set of available actions (e.g. go or no-go), and the consequences of those actions, quantified through a Net Expected Utility (NU) function.

In this chapter, we develop and apply a BDT framework designed specifically for this purpose. We first discuss the motivation and the key parameters required—including the definition of the PoM threshold ($\delta_{\%}$), the selection of the decision gate (τ^*), and the specification of the economic utility function. We then present an illustrative simulation study based on induration diameter data from a KLH challenge as a PD readout for a hypothetical immunomodulatory IMP . The simulation demonstrates how integrating scientific evidence with economic costs influences the optimal evidential threshold (τ^*) and informs efficient sample size planning (n) to maximize NU. Finally, we discuss the implications of this approach for robust, transparent decision-making in novel and uncertain scenarios typical of experimental medicine.

6.1.1 Motivating Example: KLH Challenge

In this context, we frame the PoM question as a decision problem. The primary estimand is not merely the effect size, but the probability that the drug's effect relative to placebo (or a comparator) exceeds a minimum meaningful biological effect ($\Delta_{\%} \geq \delta_{\%}$). The decision rule is to 'go' if this probability exceeds a utility-calibrated gate, τ^*.

To demonstrate this framework, we will use induration diameter elicited by KLH rechallenge (see Chapter 4) as a concrete example, although any PD biomarker could be analysed in this way. The model incorporates PoM as a latent quantity which is noisily assayed by the outcome measure.

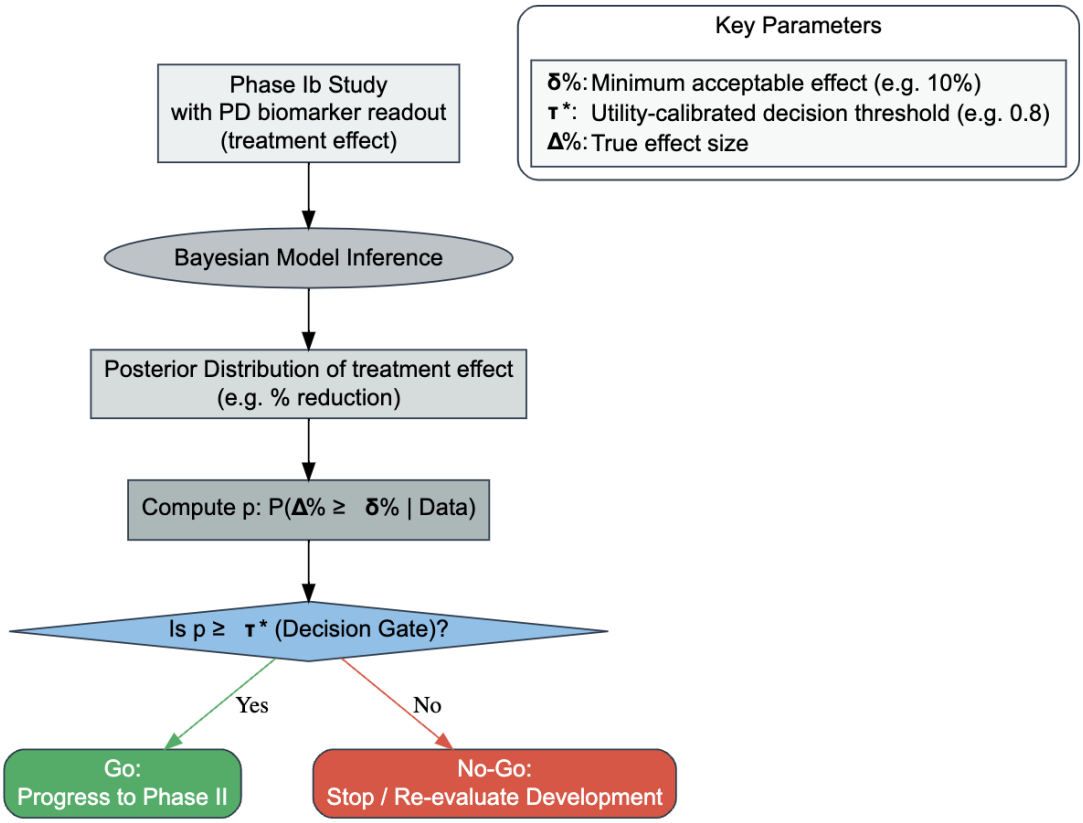

Figure 6.1: Decision approach based on hypothetical study readout of % reduction in KLH-induced skin induration in IMP treated vs. untreated participants.

6.2 Simulation Study Methods

To ensure transparency and reproducibility, the methods for the simulation study are organised according to the ADEMP (Aims, Data-generating mechanisms, Estimand, Methods, Performance Measures) framework. [132]

6.2.1 Aims ($\mathcal{A}$)

The primary aim of the simulation was to illustrate how the proposed approach may be used to derive optimal sample size (n) and decision gate (τ^*) to maximise NU for a PoM-informed Go/No-go decision. A secondary aim was to illustrate how application of conventional or informal decision criteria may lead to sub-optimal outcomes compared to this method.

6.2.2 Data-Generating Mechanisms ($\mathcal{D}$)

We consider a hypothetical phase Ib PoM study of a novel immunomodulatory drug, where PoM requires evidence of modulation of the inflammatory response to a standard 10μg KLH intradermal rechallenge (administered to all study participants), defined as the diameter (in mm) of cutaneous oedema measured 48h following rechallenge. Parameter values were based on those observed in the E_{max} dose-response model of Chapter 4.

- **Design and Outcome:** A two-arm parallel design was used (Treatment vs. Placebo). The outcome, D, was modeled as log-Normally distributed on the transformed scale $Y = \log(D + 1)$, where $Y \sim \mathcal{N}(\mu, \sigma^2)$. Because the model theoretically permits very small negative values of D, we truncated these simulated outcomes at $D = 0$ to respect the physical non-negativity constraint ($D \geq 0$).
- **Control Distribution Parameters:** The log-scale parameters were calibrated to the KLH control group posterior medians: σ^2 was set using $\sigma^{\text{true}}_{\log} = 0.597$. The control mean parameter μ_0 was derived analytically to ensure that the implied mean on the original scale matches the

target untransformed mean of $\text{mean}_{\text{ctrl}} \approx 50.3$ mm via the formula:[133]

$$\mu_0 = \log(\text{mean}_{\text{ctrl}} + 1) - \tfrac{1}{2}(\sigma_{\log}^{\text{true}})^2$$

- **True Scenarios:** Performance was assessed across $R = 1000$ replicates (per combination of n and τ^*) for true effect sizes (true % reduction in mean induration diameter $\Delta_{\%}^{\text{true}}$) used in the utility calculation:

  1. No effect (null) (0% reduction).
  2. Strong effect (40% reduction).

In addition, two additional effect size scenarios ("minimal" 10% reduction, and "moderate" 20% reduction) were included for P(Go) operating characteristics, but these were not included in the NU simulation.

6.2.3 Estimand ($\mathcal{E}$)

The primary estimand derived from the Bayesian analysis is the *Posterior Probability of PoM (p)*, which compares the estimated effect to a pre-specified 'minimum biologically significant effect' threshold, set to $\delta_{\%} = 10\%$:

$$p = P(\Delta_{\%} \geq \delta_{\%} \mid \text{data}).$$

6.2.4 Methods ($\mathcal{M}$)

Simulations were performed using R (version 4.4.0) in RStudio (2024.04.1+748). Bayesian models were fit using Stan via the `cmdstanr` interface. [134]

- **Model Specification:** $Y_i \sim \mathcal{N}(\alpha + \theta T_i, \sigma^2)$, where $Y_i = \log(D_i + 1)$ and $T_i \in \{0, 1\}$ indicating placebo and treatment respectively.
- **Priors:** Priors were set to be weakly informative:
  - $\alpha \sim \mathcal{N}(\mu_0, 0.25^2)$ (calibrated location).

- $\theta \sim \mathcal{N}(0, 0.5^2)$ (weakly-informative prior).
- $\sigma \sim$ truncated-$\mathcal{N}(0.55, 0.15^2)$ (calibrated to observed KLH variability, Chapter 4).

- **Computational Engine and Diagnostics:** Posterior inference was performed using the Pathfinder algorithm implemented in Stan, which provides a rapid approximate posterior draw set via a parallel quasi-Newton variational approach.[135]

- **Optimization Strategy and Decision Rule:** The optimal design (τ^*, n) was identified via a grid search, maximizing the NU over $n \in \{6, 8, \ldots, 28\}$ and $\tau^* \in \{0.60, 0.65, \ldots, 0.95\}$. The decision rule is to declare Go if $p \geq \tau^*$.

6.2.5 Performance Measures ($\mathcal{P}$)

- **Primary Measure: Net Expected Utility (NU):** To derive NU, the immediate output of the PoM study (the binary decision to 'Go' or 'No-go') was mapped to long-term economic value via a translational probability model.

  The model assumes that observing a valid PoM signal enriches the probability of success in the subsequent Phase II trial. Based on pipeline attrition data [10], we assigned a conditional probability of phase II success of $\pi_{\text{pos}} = 30\%$ if PoM is established, compared to a baseline success rate of $\pi_{\text{neg}} = 5\%$ if the drug progresses without PoM.

  The Expected Value of Success (E[VS]) is calculated as the expected downstream value from making a 'Go' decision under a simplified binary 'state of the world': where the drug is either truly effective (H_1, prior probability $w_1 = 0.2$, represented by the "strong" 40% reduction scenario) or ineffective (H_0, prior probability $w_0 = 0.8$, represented by the "null" 0% reduction scenario). Specifically:

$$E[\text{VS}] = V_{\text{succ}} \times [w_1 \cdot P(\text{Go}|H_1) \cdot \pi_{\text{pos}} + w_0 \cdot P(\text{Go}|H_0) \cdot \pi_{\text{neg}}] \tag{6.1}$$

  where V_{succ} represents the Net Present Value of an IMP following a successful phase II study result.

The overall expected financial return from the decision, based on the expected value of success versus the costs of progression, and phase I study design (determined by per arm sample size):

$$\mathrm{NU}(n) = E[\mathrm{VS}] - E[\mathrm{C_{prog}}] - \mathrm{C_{design}}(n)$$

- **Input Parameters:** The specific economic and probability values used in the simulation are detailed in Table 6.1.

Table 6.1: Economic and Probabilistic Parameters for the NU Function. Values are illustrative, scaled to arbitrary units where 100 represents the value of a phase II success.

Parameter	**Symbol**	**Value**	**Description**
Economic Utilities			
Value of Success	V_{succ}	100	Net Present Value (NPV) if the drug succeeds in phase II.
Cost of Phase II	C_{prog}	10	Fixed cost incurred upon a 'Go' decision (e.g., running the subsequent trial).
Cost per Subject	c_{pp}	0.01	Marginal cost per Phase I participant.
Translational Probabilities			
Prior: Effective	w_1	0.20	Prior probability that the drug mechanism is truly effective (H_1).
Prior: Null	w_0	0.80	Prior probability that the drug is ineffective (H_0).
PTS (with PoM)	π_{pos}	0.30	Prob. of Phase II success given PoM is declared (Go).
PTS (no PoM)	π_{neg}	0.05	Prob. of Phase II success given PoM is not declared (or False Go).

- **Other Cost Components:**
  - $\mathbf{E}[\mathrm{C_{prog}}]$ (Expected Cost of Progression): The cost incurred if a Go decision is made, weighted by the probability of a Go decision.
  - $\mathrm{C_{design}}(\mathbf{n})$: The direct cost of the Phase I study, defined as a function of sample size: $2n\, c_{\text{pp}}$.

- **Operating Characteristics:** The long-run operating characteristics of the decision rule were assessed via the Probability of Go ($P(\text{Go}) = \Pr[p \geq \tau^*]$).

6.3 Results

6.3.1 Probability of a Go Decision

The operating characteristics of the decision rule are displayed in Figure 6.4. As expected, the probability of a 'Go' decision ($P(\text{Go})$) reflects the interaction of the true effect size (Δ) and decision threshold τ^*. Notably, the probability of a Go decision is low at classical probability thresholds (e.g. 0.9 or 0.95), especially in marginal effect scenarios (e.g. 10% true effect size) but also to a lesser degree where the true effect is substantially above (e.g. 40%). Notably, as sample size increases the separation of effect curves becomes more pronounced, more effectively separating true signal from statistical noise, especially at more moderate evidential thresholds (e.g. τ^* 0.65-0.8).

6.3.2 Net Utility According to Decision Gate and Sample Size

The relationship between design parameters and NU is shown in Figure 6.5. This figure highlights the consequences of varying sample size and τ^*, in particular the marginal effects of progressively more stringent decision thresholds (panel A) and increasing sample sizes (panel B). In this scenario, it can be seen that NU increases monotonically with increasing sample size but approaches an asymptote and effectively plateaus around n=18 (panel B). At sample sizes of n=18 or above, an optimal decision threshold is observed at $\tau^* = 0.80$ but is relatively flat between 0.7-0.85 (panel c), and declines more sharply at higher thresholds, thus defining optimised parameters for the hypothetical PoM study.

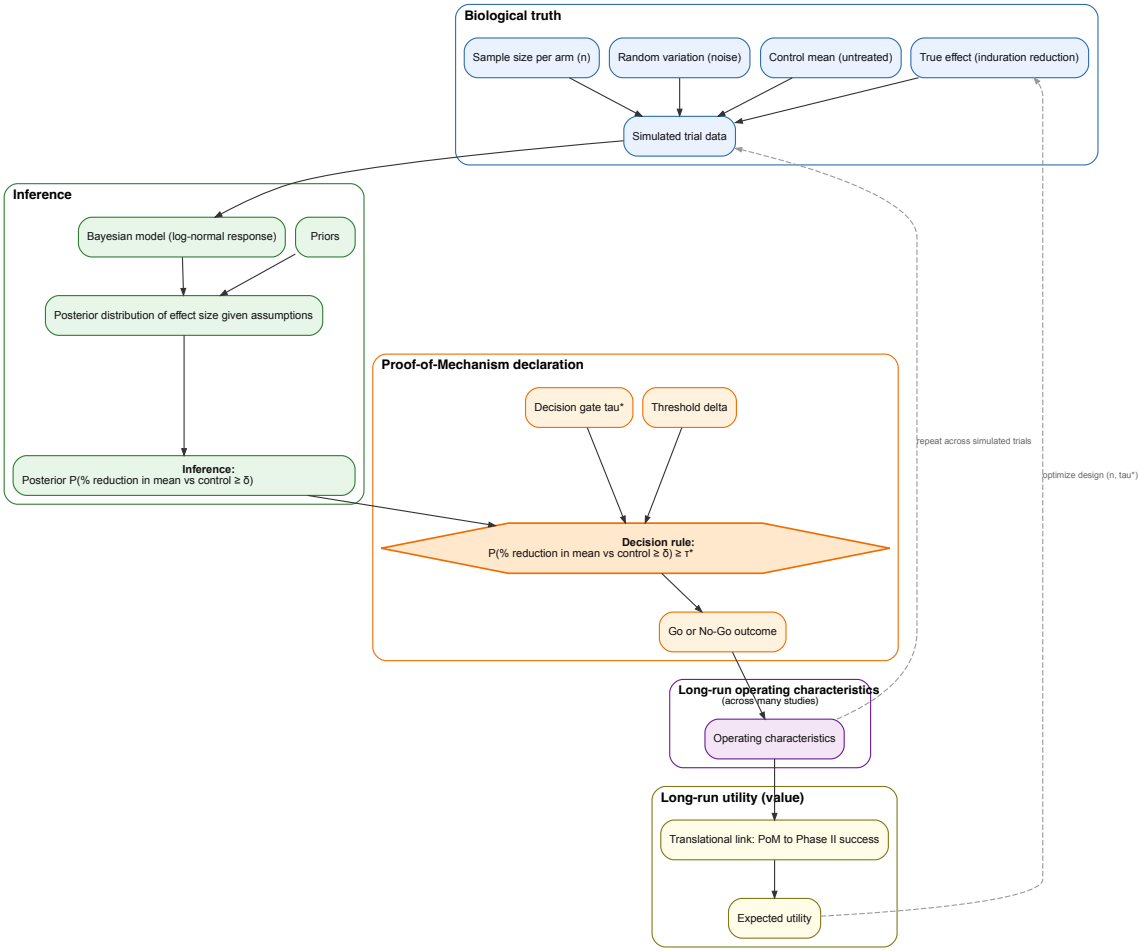

Figure 6.2: Structure of the Proof-of-Mechanism decision simulation. Each simulated study samples data under an assumed *biological truth* (true effect, control mean, random variation), and sample size per arm n, and is analysed with a Bayesian model to obtain a posterior distribution of the percent reduction in mean versus control. From this posterior, the probability that the reduction exceeds a pre-specified threshold δ is computed (*inference*), and a *Proof-of-Mechanism* is declared if this posterior probability is greater than or equal to the decision gate τ^*, leading to a *Go/No-go* decision. Repeating this process across many simulated studies yields *long-run operating characteristics* (proportion of Go vs No-go decisions), which are then linked through a translational model based on estimated success probabilities conditional PoM declaration, to estimate *long-run expected utility* (the expected value of acting on the decision). The framework can be used to optimise design parameters such as n or τ^*, and assess value of model optimisation (predictive variables).

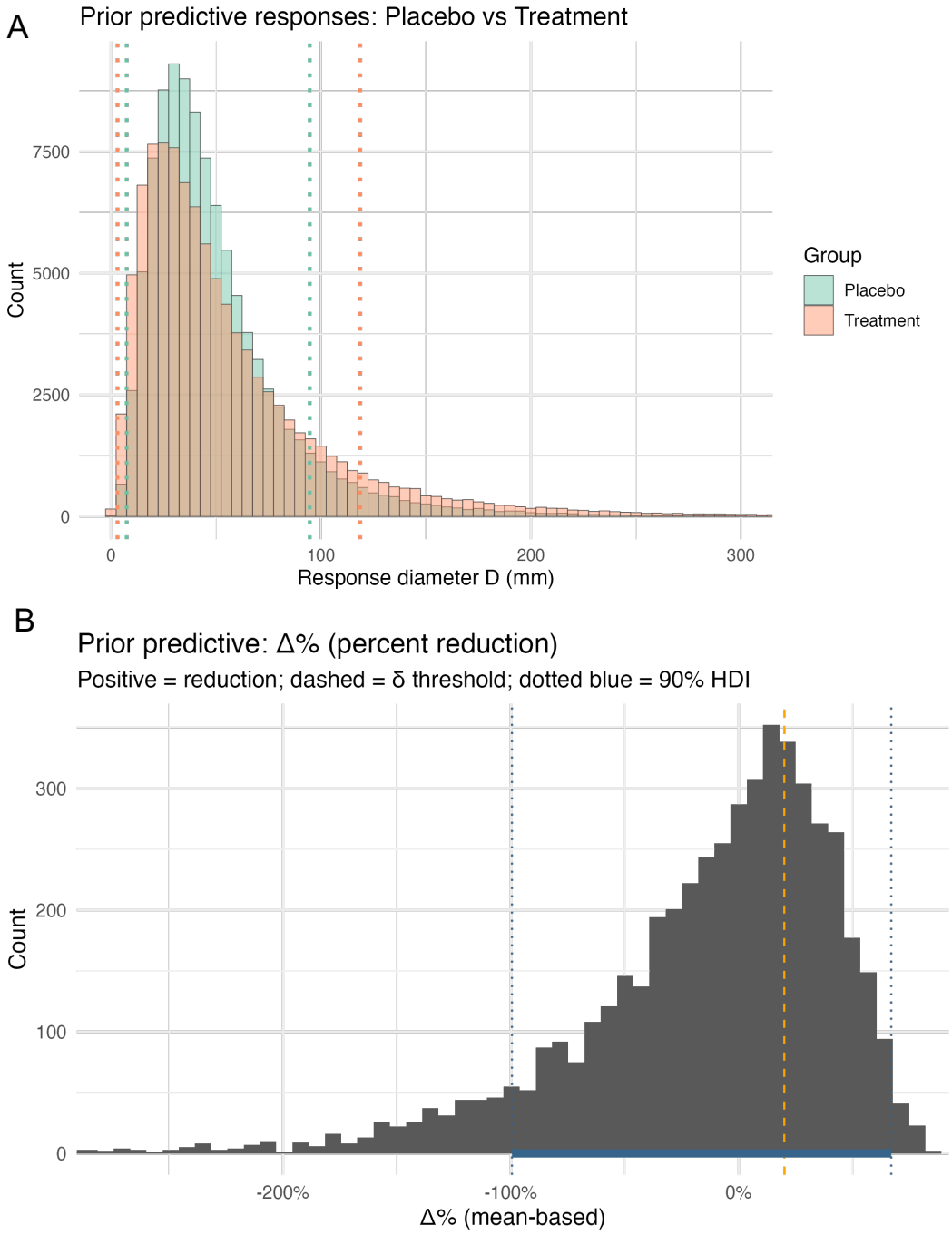

Figure 6.3: Prior distributions for the key parameters of the model, which analyses the transformed outcome log(Diameter + 1). The Untreated ('placebo') Mean Parameter (α) is given a weakly informative Normal prior, $\mathcal{N}(\alpha_0, 0.25^2)$, with its location α_0 calibrated using the median response from the KLH control data. The Treatment Effect (θ) receives a weakly-informative Normal prior, $\mathcal{N}(0, 0.5^2)$, reflecting no initial expectation of effect. The Log-Scale Standard Deviation (σ) is assigned a truncated Normal prior, $\mathcal{N}(0.55, 0.15^2)$, with its centre aligned near the marginal log-scale variability estimated from the same KLH control distribution.

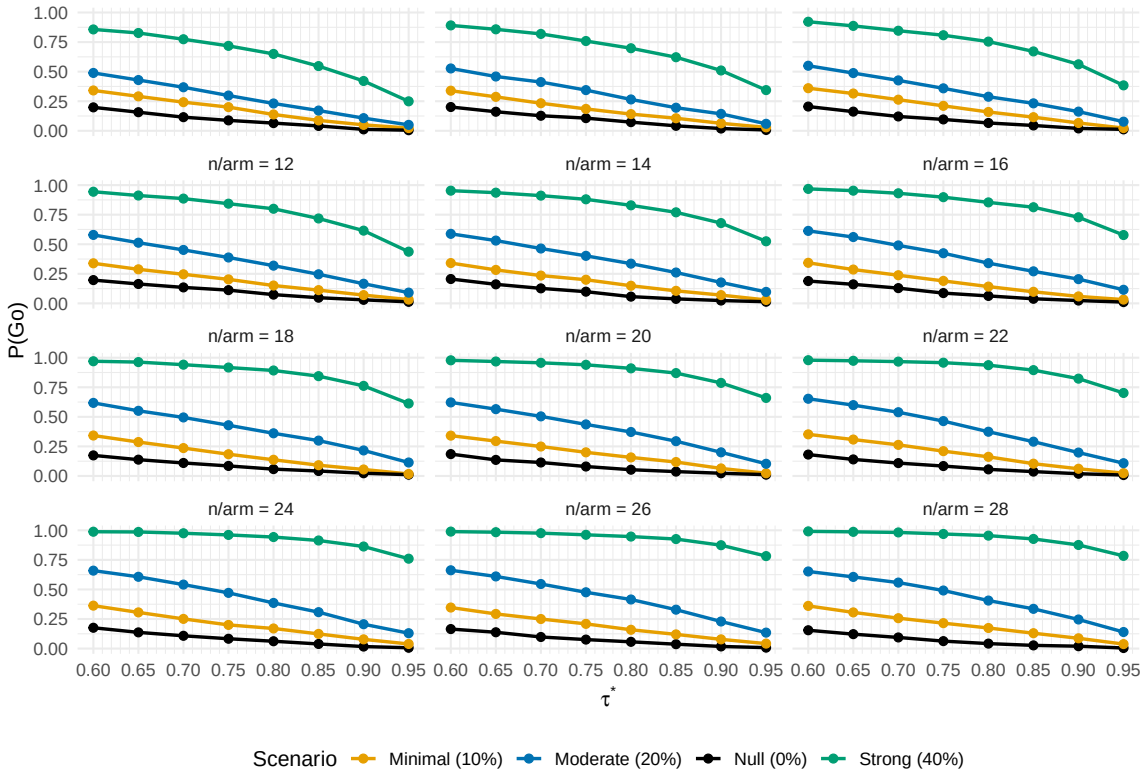

Figure 6.4: Probability of a 'Go' Decision (P(Go)) vs. Decision Threshold (τ^*), Faceted by sample size. P(Go) is the long-run probability that the posterior probability of achieving the minimum acceptable effect ($\Delta_{\%} \geq \delta_{\%} = 10\%$) is greater than or equal to the decision gate τ^*. P(Go) (y-axis) is shown as a function of the decision threshold τ^* (x-axis) across different true effect scenarios. Each panel represents a different sample size per arm (n).

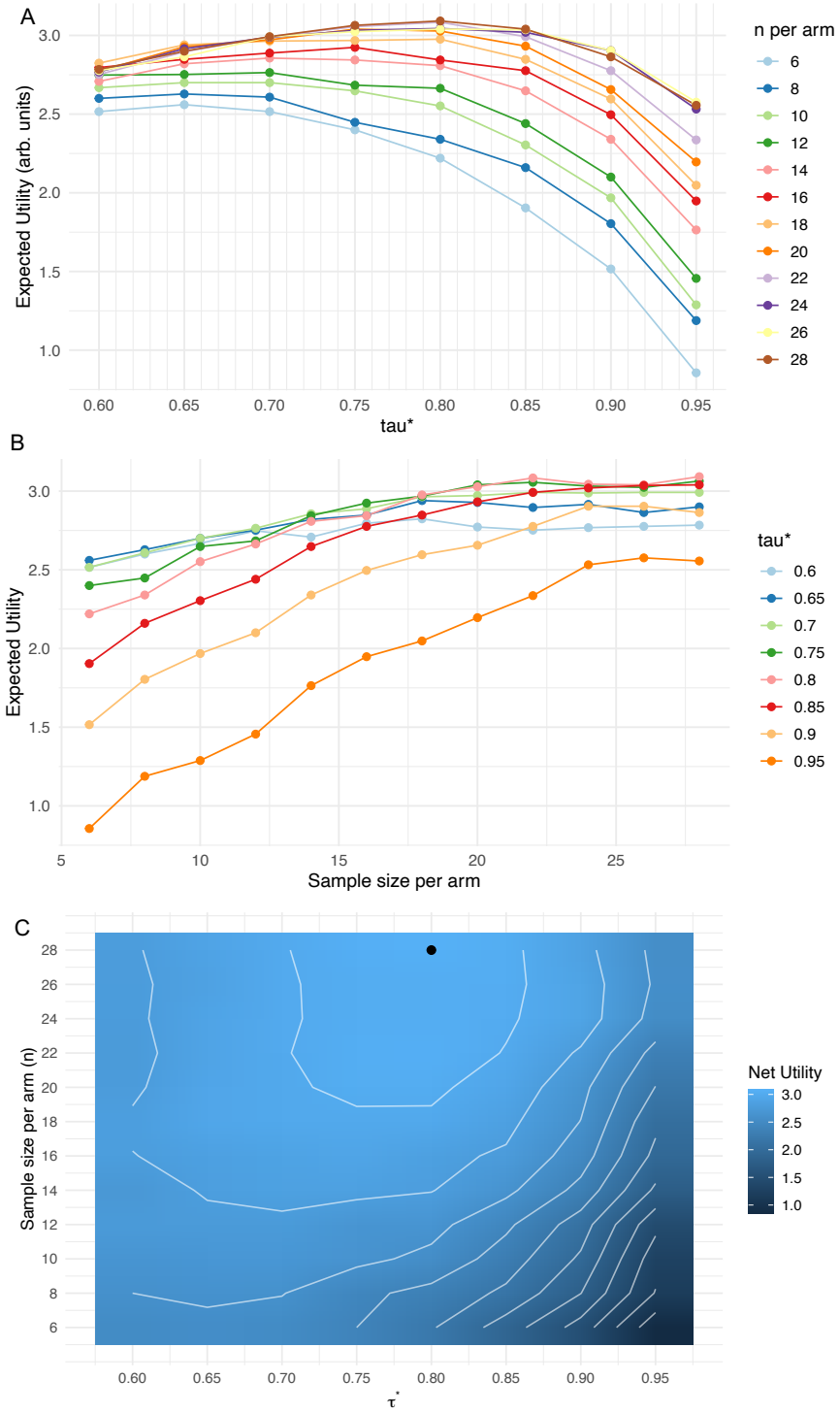

Figure 6.5: Net expected utility optimization. Trade-off between the design parameters (τ^* and n) and the resulting Net Expected Utility (NU(n)), which combines the value of successful progression with associated costs (design, progression). A: NU vs. Decision Threshold (τ^*): NU (y-axis) as a function of the probabilistic decision gate τ^* (x-axis), with a separate line for each investigated sample size per arm (n). B: Expected Utility vs. Sample Size (n): NU (y-axis) as a function of the sample size per arm n (x-axis), with a separate line for each decision gate τ^* investigated. C: 2D contour plot τ^* (x-axis, n (y-axis) and NU (coloured contours). The global optimal design maximizes NU across the full parameter grid, identifying the most economically efficient combination of n and τ^*, given the model assumptions.

6.4 Discussion

6.4.1 Key Findings

In this work, we have illustrated how the quantification and propagation of uncertainty through to a utility-based endpoint offers a transparent probabilistic framework for early-phase decision-making. Based on the specific assumptions regarding the performance of the PoM outcome measure (% reduction in KLH-challenge-induced skin induration in treated versus placebo), biological significance threshold, true underlying effects, conditional translational success probabilities, and utility, a utility informed study design could identify as optimal a PoM study of $n = 18$ per arm, with a decision rule $P(\%\text{ reduction in mean versus control} \geq \delta) \geq 0.80$. In this simulation, this logic would therefore argue for a larger phase I PoM study, with lower evidential thresholds than may be conventionally employed in contemporary clinical trials.

A key finding is that standard frequentist significance thresholds (e.g., $\alpha = 0.05$) may be markedly suboptimal when analysing HIC-derived pharmacodynamic biomarker studies, a finding significantly influenced by the asymmetric utilities associated with false-stop versus false-go decisions in early development.

6.4.2 Comparison with Prior Literature

While decision-theoretic approaches have been applied in Phase II and III trial designs, a review of the literature identified a paucity of published applications to Phase I decision-making.[136–138] This gap likely reflects the historical separation between the analysis of biological evidence and the evaluation of program viability, and the absence of suitable pharmacodynamic biomarkers for PoM interrogation (such as those elicited by HIC). Standard early-phase methodology predominantly focuses on operating characteristics defined by frequentist error rates (e.g., toxicity stopping rules or hypothesis tests), while program-level decision-making often relies on separate economic models that utilise static assumptions of success.[139] Consequently, there is a lack of integrated frameworks that utilise the full Bayesian posterior distribution of the biomarker response to dynamically inform the

'Go' decision. The framework presented here addresses this gap by creating a direct translational link between the biological signal and the decision context, ensuring that the uncertainty inherent in the HIC model is explicitly accounted for in the progression criteria.

6.4.3 Strengths and limitations

A key strength of this study is the integrated calibration of critical study design decisions (n and τ^*) in a manner directly aligned with the objective of the development program. The strength of the BDT approach outlined here is that it enables the assignment of appropriate weight to prior assumptions, commensurate with the degree of scientific certainty. This is particularly critical in the context of immune challenge studies, which, by definition, serve as uncertain representations of the pathological processes targeted by the drug. This framework could be used to inform the design of early-phase study where a PD readout can be plausibly simulated and its link with late phase success estimated.

However, the results presented here should be interpreted as an illustrative rather than an exhaustive framework for quantitative Phase Ib PoM decision-making, and several limitations should be noted.

Firstly, the simulation assumes a simplified, linear utility function. In this formulation, utility is defined in terms of captured value arising from progression decisions, ignoring complexities such as opportunity costs and second order effects of success or failure. In practice, pharmaceutical decision-making is highly non-linear, involving complex Real Options-based considerations, such as the flexibility to expand into multiple indications, the option to delay investment, and portfolio-level resource constraints. While our linear model provides a necessary first step in linking biomarker data to utility, it does not fully capture the dynamic nature of managing a clinical portfolio. Furthermore, the cost of design was modelled as a linear function of the number of participants. In practice, it may be more realistic to more heavily penalise larger studies to reflect the substantial additional complexities inherent in an atypically large Phase I trial, such as the cost of manufacturing sufficient IMP, and the time-cost of executing larger studies. In this setting, rather than approaching a

plateau, the utility curve would be expected to decline in the setting of larger sample sizes, rather than approaching an asymptote as seen in this simulation.

Secondly, the translational assumptions, such as probability of success given PoM, and the true state of the world (probability of true efficacy of a drug entering phase I) are represented as low resolution, binary states. While conceptually and computationally convenient this represents a gross oversimplification of reality. It would clearly be desirable to precisely inform these assumptions using real-world data relevant to the IMP in question. However this is likely to remain challenging, especially for novel mechanisms. That said, as Scannell and colleagues have argued, even modest improvements in predictive validity of a decision tool may yield substantial benefits in overall pharmaceutical R&D productivity.[140]

6.5 Conclusion and Future Directions

The work presented here provides a basis for further methodological development and empirical validation of the proposed decision-theoretic framework. Future research should focus both on extending the formal structure of the model and on strengthening the translational grounding of its key assumptions.

First, the framework may be extended beyond simple two-arm parallel-group designs to accommodate dose–response settings, active comparators, and multi-arm decision problems. In many realistic development scenarios, the relevant choice is not a binary "Go/No-go" decision, but the selection among multiple candidate regimens or the decision to expand or curtail a cohort. Integration of concepts from Value of Information analysis, including the Expected Value of Sample Information (EVSI), would provide a principled basis for determining when additional data collection is expected to yield sufficient decision value to justify further investment.[141]

Second, broader integration of pharmacodynamic (PD) readouts should be explored, including composite and multivariate biomarkers derived from multiple tissue- and blood-based endpoints. Systematic quantification of biomarker operating characteristics—such as the magnitude and sources of heterogeneity, and sensitivity to modulation—would enable more informed calibration of study

design parameters and reduce uncertainty in the resulting utility estimates. This could include formal evaluation of explanatory covariates (e.g. baseline HIC response, enabling a within-subject experimental design) and benchmarking of conditional success probabilities against historical data from mechanistically related immunomodulatory agents.

Third, latent-variable approaches based on validated Quantitative Systems Pharmacology (QSP) models warrant further investigation. Such models may offer a closer correspondence between observed biomarker perturbations and hypothesised clinical mechanisms than single observable endpoints, potentially improving the interpretability and comparability of PoM assessments across indications and target classes.

Finally, given the decision-critical role of the simulation framework, structural assumptions should be sufficiently rich to reflect real-world development constraints and incentives.[142, 143] These include non-linear cost structures, opportunity costs, and portfolio-level resource allocation. At the same time, increased model complexity must be balanced against interpretability. Prior elicitation frameworks such as Sheffield Elicitation Framework (SHELF) may support formal encoding of subjective assumptions,[144] but parallel effort is required to ensure that model structure and implications can be communicated transparently to non-statistical stakeholders. Approaches such as Goal Structuring Notation, originally developed for high-consequence engineering systems and adapted for biological modelling, offer one possible avenue for making decision logic explicit and auditable.[145]

In summary, this chapter advances a formal approach in which PD biomarkers are treated not primarily as objects of hypothesis testing, but as inputs to explicit decision rules calibrated by their downstream consequences. The principal conceptual contribution is the integration of Bayesian uncertainty quantification with programme-level utility, enabling transparent optimisation of design parameters such as sample size and evidential thresholds. While the present implementation is intentionally simplified, it provides a foundation for more empirically grounded and operationally realistic applications in early-phase immunomodulatory drug development.

Chapter 7

Discussion

Chapter summary

In this final chapter, I summarise the key findings from each of the five analytical thesis chapters and highlight their unique contribution to the literature. I highlight the main limitations of these analyses, and drawing on the key findings, I outline suggested directions for future research and application of HIC paradigms in early phase drug development.

7.1 Overview

The central premise of this thesis is that the productivity crisis in drug development is substantially driven not by a lack of potential druggable targets or novel molecules, but by lack of effective strategies to validate them early in the clinical pathway. As outlined in Chapter 1, the inability to demonstrate PoM in Phase I leads to costly attrition in Phase II. The overarching aim of this research was to further enhance experimental and analytical platforms for HIC in drug development, which necessitates a transformation from niche experimental medicine approach into a robust decision-orientated toolkit. Acknowledging that this task is substantially larger than a single DPhil project, I focused on high priority, tractable objectives: characterising and optimising key operating characteristics in two potential challenge models (Chapters 2 4), systematically reviewing existing evidence to identify areas for model optimisation (Chapter 3), attempting to identify an improved method for lymph node sampling (thereby improving acquisition of tissue-based immune challenge

readouts, Chapter 5), and finally integrating PD biomarker data derived from immune challenge into a quantitative decision framework (Chapter 6). Collectively, this body of work demonstrates that while HIC models are subject to biological heterogeneity, this variability is characterisable and manageable. When coupled with optimised tissue-based endpoints and BDT, HIC paradigms offer a feasible, ethically sound, scalable paradigm to address the crisis of productivity in immunomodulatory pharmaceutical R&D.

7.2 Summary of Thesis Objectives and Key Findings

7.2.1 Findings from the observational immunological study of responses to MVA-BN vaccination

The objective of this chapter was: *To characterise the immune response to a marketed prophylactic vaccine (MVA-BN) in a healthy cohort, to identify predictors of response, and to assess its potential value as a viral HIC agent for fundamental immunology and for PD evaluation in immunomodulatory drug development.*

To characterise a viral-vector HIC paradigm, I analysed the immune response to MVA-BN in a prospective cohort of healthy volunteers ($n = 34$) during the 2022 mpox outbreak. The study demonstrated that MVA-BN is a potent inducer of CD8+ T-cell immunity, with a 92% response rate at Day 14 and a response magnitude significantly biased towards CD8+ rather than CD4+ compartments (median 245 vs 124 SFC/10^6 PBMC). However, the data revealed profound biological heterogeneity, with variable serological and (systemic) cellular immunogenicity (e.g. peak CD8+ responses ranging from 0 to 897 SFC/10^6 PBMC). Crucially, while peripheral blood transcriptomic signatures could predict the subsequent antibody response, they failed to predict this T-cell heterogeneity. This study confirmed the potential of MVA-BN as a novel HIC, but highlighted the need for further optimisation and characterisation, particularly the necessity and value of extending the model to include readouts based on skin and lymph node sampling. Vaccine supply issues, and

the potential for pre-existing immunity related to prior vaccination and (rare) exposure to mpox, suggested the value to focus on a alternative challenge agent with established HIC application: KLH.

7.2.2 Findings from the systematic review published KLH challenge studies

The objective of this chapter was: *To systematically identify, collate, and synthesise HIC studies employing KLH, in order to define design limitations and inform a research agenda for optimisation of KLH HIC paradigms in drug development.*

To evaluate the current state of the literature, I conducted a systematic review of KLH challenge studies published between 1994 and 2022 (n = 46), specifically focusing on their application in early-phase drug development. The review identified a clear trend: the paradigm is transitioning from a seldom-used tool for immunotoxicity assessment and experimental medicine to a distinct platform for interrogation of IMPs PD, particularly those targeting T-cell co-stimulation (e.g., CD28, OX40, CD40 pathways). However, the review revealed marked methodological heterogeneity that currently limits the platform's deployment across drug development programmes. Primary challenge doses varied 50-fold (range 100 µg to 5000 µg), and reporting quality was frequently poor when assessed against the TIDieR checklist, with many studies failing to specify the exact formulation or source of the antigen. The majority of studies (93%) relied solely on systemic antibody titers (IgG/IgM) as the primary endpoint–a striking finding given the availability of modern immunological methods arguably better suited to focused assessment of immunological pathway activity. While sufficient for detecting gross immunotoxicity, the reliance on peripheral humoral readouts represents a missed opportunity for PD assessment at high mechanistic resolution in drugs such as those targeting T cell co-stimulation. These findings confirmed that before KLH could be utilised as a decision-making tool, further study was required to define the dose-response relationship of antigen rechallenge and characterise tissue-level endpoints over a timecourse—the primary objective of the subsequent experimental work.

7.2.3 Findings from the KLH challenge study

The objectives of this chapter were:

To further characterise and optimise the KLH immune challenge model to advance its suitability for early-phase Proof-of-Mechanism studies in healthy volunteers:

a. *To determine the effect of co-administered adjuvants (Alhydrogel and Montanide ISA-51) on the magnitude and $CD4^+$ T_h1 / T_h2 polarisation of the resultant systemic primary immune response to KLH in healthy volunteers.*

b. *To quantify the effect of KLH rechallenge dose and identify an optimal dose for future PoM studies.*

c. *To characterise the temporal evolution of cutaneous immune responses to KLH rechallenge using multi-parameter flow cytometry, and thereby identify timepoints optimally suited to interrogate mechanistically distinct immune processes.*

This study systematically characterised the KLH challenge model to establish critical parameters for its use in PoM studies. A key finding was the dissociation between humoral and cellular immunogenicity driven by adjuvant selection; while both Montanide ISA-51 and Alhydrogel elicited comparable antibody titres, Montanide ISA-51 proved significantly superior in inducing systemic T cell responses. Consequently, Montanide may be considered the optimal adjuvant for studies requiring a robust peripheral $CD4^+$ T cell signal. Furthermore, Bayesian Emax modelling identified 10 μg as the optimal intradermal rechallenge dose for this primary immunisation regimen. This dose balances detectability with tolerability and sensitivity, achieving a high probability of response without the saturation and non-specific myeloid dominance observed at higher doses.

The investigation also redefined the optimal timing and method of outcome assessment. Contrary to the reliance on the standard 48-hour clinical endpoint, we observed a dynamic temporal shift in the cutaneous infiltrate: early responses (48 h) were characterised by a mixed myeloid/lymphoid population with balanced Th1/Th2 phenotypes, whereas later timepoints (Days 5–14) revealed a lymphocyte-predominant infiltrate with a progressive Th1 skew and resident memory phenotype,

highlighting distinct kinetics across key cell populations. This dictates that biopsy timing must be mechanistically tailored to the investigational drug and IMP target. Finally, comparison of assessment modalities confirmed that simple clinical measurement of induration remains the most robust non-invasive predictor of the underlying cellular infiltrate, outperforming objective imaging metrics which serve primarily as quality assurance tools rather than surrogates for adaptive immune magnitude.

7.2.4 Findings from the lymph node contrast-enhanced ultrasound experimental medicine study

The objective of this chapter was:

To evaluate, using an experimental medicine approach, the feasibility of contrast-enhanced ultrasound for identifying primary draining lymph nodes in healthy volunteers to enable tissue-based sampling in HIC and PoM studies.

This pilot study applied an experimental medicine approach to evaluate Contrast-Enhanced Ultrasound (CEUS) as a tool for precision lymph node targeting. While the clinical endpoint was not met, the study validated the utility of adaptive trial designs in early-phase procedural development. The primary finding was the lack of utility for CEUS in this context; the technique failed to reliably identify primary draining lymph nodes in healthy volunteers, with no contrast enhancement observed within discrete lymph node architecture in any participant (0/4). Consequently, CEUS does not offer a tangible advantage over standard anatomical ultrasound for guiding invasive sampling in this model, although given this finding was at odds with the successful results reported for sentinel lymph node identification in breast malignancy, there remains the potential for future studies to derive an effective imaging method.

Despite the negative primary outcome, the study highlighted the advantages of a Bayesian adaptive sample size design. By continuously updating the posterior probability of success, the adaptive design correctly identified the futility of the intervention at the earliest possible stage ($n = 4$). This triggered an early study termination, preventing unnecessary procedures and conserving

resources.

7.2.5 Findings from the simulation study

The objective of this chapter was:

To demonstrate how immune challenge-derived PD readouts may be formally integrated within a Bayesian decision framework, and to examine how utility-calibrated decision rules can optimise Go/No-Go thresholds relative to conventional Phase I decision strategies.

This study demonstrates that integrating PD data into a BDT framework allows for the optimisation of study design parameters based on economic utility rather than arbitrary statistical convention combined with categorical heuristics. Specifically, the simulation identified a plateau in maximal net expected utility at a sample size of $n = 18$ or greater per arm, with a decision gate (probability of a meaningful PD effect given the data) of $\tau^* = 0.80$. While early-phase groups are typically smaller, the analysis showed that increasing the sample size modestly provides the necessary statistical power to distinguish meaningful biological signals from the inherent variability of the KLH model, with economic returns plateauing only at this higher number. Furthermore, the results highlight that standard frequentist significance thresholds (e.g., $\alpha = 0.05$) are likely economically suboptimal in the PoM context. Such stringent criteria disproportionately increase the rate of "False Stops", and thereby failing to realise the disproportionate economic gains associated with attaining marketing authorisation of an effective drug. While these findings should be interpreted as illustrative–as the specific optimal parameters were derived from a intentionally simplified model–the simulation represents proof-of-concept for a quantitative decision making framework worthy of further study and eventual application in drug development settings. By formally quantifying sources of uncertainty and calibrating these to economic outcomes, the proposed framework establishes a rational and transparent decision-making process that balances the informative value of translational biomarkers against the inherent uncertainties of early-phase development.

7.3 Integration with Existing Literature

This thesis advances the existing literature by bridging the disconnect between the biological potential of HIC models and their practical application as decision tools for drug development. It builds upon a substantial foundation of prior experimental work, acknowledging that while previous iterations of the KLH model successfully demonstrated the feasibility of inducing neoantigen responses in humans, their translational utility was often constrained in the literature by methodological heterogeneity and a reliance on systemic readouts. This research addresses some of these historical limitations by systematically defining the challenge parameters—specifically regarding adjuvant selection, dose-response relationships, and tissue sampling—thereby providing the standardisation required to align the KLH model with contemporary requirements.

Concurrently, this work integrates with and expands upon the conceptual frameworks defining modern drug development, specifically the industry-wide shift towards establishing PoM in early-phase trials. While the theoretical value of PoM is well-supported by retrospective analyses of portfolio success rates, the mechanism for prospectively integrating these biological signals into decision-making has largely remained qualitative and heuristic in the literature. By formalising PoM as a quantitative Bayesian decision problem, this thesis provides a novel analytical methodology to operationalise these concepts. Ultimately, the main contribution of this research to the field lies in the synthesis of these experimental and analytical domains; it moves the literature beyond the question of whether immune challenges can be performed, to an evidence-based demonstration of how they can be utilised as robust, decision-enabling platforms.

7.4 Strengths

The primary strength of this research lies in the breadth and integration of its methodological framework. Within the resource and temporal constraints of a doctoral program, this thesis executed a research arc comprising a systematic literature review, an observational immunological cohort study, three distinct interventional experimental medicine studies, and a statistical simulation. This multi-

modal approach permitted a comprehensive evaluation of the HIC platform, allowing the research to address the central problem of early-phase attrition from multiple vantage points simultaneously: historically (via systematic review), biologically (via the MVA-BN and KLH studies), technically (via the contrast-enhanced ultrasound study), and analytically (via the Bayesian decision simulation).

Furthermore, the work is strengthened by the direct, data-driven linkage between problem identification and experimental execution. Rather than applying a pre-existing protocol, the systematic review (Chapter 3) was first utilised to diagnose the specific limitations of the existing literature—specifically, the lack of standardised dosing and the reliance on serological endpoints. These insights directly informed the design of the subsequent KLH optimisation studies (Chapter 4), ensuring that the experimental optimisation was targeted precisely at the areas of highest uncertainty.

Finally, the thesis moves beyond the standard scope of experimental medicine by formally integrating biological data with decision theory. While many studies successfully characterise immune responses, few attempt to quantify the value of that information in a drug development context. By coupling the biological optimisation of the KLH model with a Proof-of-Concept Bayesian simulation (Chapter 6), this research provides not just a refined experimental tool, but a translational framework. This demonstrates that the generation of complex PD biomarkers is not only scientifically feasible, but, when interpreted through a utility-based framework, economically rational for high-stakes decision-making.

7.5 Limitations

Several limitations of this work should be acknowledged. First, consistent with many experimental medicine protocols, the sample sizes in the KLH challenge studies reported were small. While sufficient to capture strong signals (e.g. effects of adjuvants in the KLH study, qualitative time course of major immune cell populations in the skin post KLH rechallenge), larger studies are necessary to adequately characterise the statistical operating characteristics of specific biomarkers, including the determinants of the variability observed in key immune parameters.

Second, the inability to reliably identify lymph nodes via contrast-enhanced ultrasound (Chapter

5) highlights the technical and logistical challenges inherent in implementing complex experimental medicine protocols, with time and resource constraints preventing further attempts to derive a working imaging approach (or conclusively confirm its futility).

Third, the analysis of Modified Vaccinia Ankara - Bavarian Nordic (MVA-BN) responses (Chapter 2) was observational and necessarily limited in scope, for example precluding skin or lymph node sampling and resultant insights–this should be addressed in future studies.

Finally, a fundamental epistemic limitation of this work must be acknowledged. While mechanistically rational and supported by observational evidence[10], successful application of HIC enabled PoM studies is predicated on the assumption that effects observed in health volunteer populations can probabilistically inform subsequent clinical efficacy, e.g. the premise that inhibiting a Delayed-Type Hypersensitivity (DTH) response to KLH has some predictable relationship clinical efficacy in conditions such as Rheumatoid Arthritis or Inflammatory Bowel Disease for a specific drug. Such an assumption must be critically evaluated for each individual case, ideally with ample empirical evidence and validation. Prospective validation of specific HIC paradigms is operationally and financially challenging, and may always lag the leading edge of novel immunomodulatory drug development.

7.6 Implications for Drug Development

The findings presented here offer immediate practical applications for the design of PoM studies. Strategically, HIC studies are best positioned within Phase Ib, serving as a bridge between First-in-Human safety trials and Phase II efficacy studies. Regarding agent selection, KLH remains well-suited for interrogating $CD4^+$ T-cell dependent mechanisms, whereas alternative antigens such as MVA-BN may offer a viable alternative for innate or viral response pathways, subject to further optimisation and characterisation. The results of the simulation study in Chapter 6 argue for an integrated decision making approach grounded in BDT principles. As illustrated, optimisation of sample sizes and decision thresholds can yield disproportionate gains in Net Expected Utility by enabling the integration of PD biomarkers while simultaneously mitigating the risk of incorrect

stopping decisions.

Furthermore, by illustrating the economic value of"better information", the BDT approach highlights the value in investment in enabling studies to develop novel or enhanced HIC approaches suitably tailored to the specific needs of the drug development program. Partnerships between industry and academia represent a path forward in this regard, with academic groups contributing to methodological development and integration of cutting-edge immunological approaches, supported by industrial partners with deep understanding of commercial realities and operational exigencies. The resultant methodological innovation derived from such partnership is likely to have wider benefit, given the impacts on R&D efficiency, in addition to a generation of platforms capable of answering questions of direct academic interest.

7.7 Conclusions and Future Directions

This thesis demonstrates that HIC models represent feasible and informative tools for drug development rather than mere academic exercises. By characterising biological variability and integrating it within a statistical framework that accounts for uncertainty, HIC can be transformed from an exploratory endpoint into a robust decision-enabling toolkit. When conducted with appropriate rigour, these models offer a solution to the high attrition rates seen in Phase II trials. They permit investigators to address the most critical question in drug development—whether a molecule exerts the hypothesised effect within a human system—at a stage where the cost of investigation is manageable and the value of the answer is maximal.

To transition these findings from academic proof-of-principle to a scalable industrial platform, future work must address systemic bottlenecks at the interface of biology, statistics, and regulation. The field must adopt a regulatory science perspective, defining a clear Context of Use and establishing a stepwise validation path comprising analytical validity, clinical validity, and translational value. To support a broader range of therapeutic indications, the challenge repertoire requires expansion to develop a "library" of agents targeting distinct immunological and inflammatory pathways. Additionally, a critical validation step involves pharmacological benchmarking, wherein HIC responses

are tested against approved therapies to calibrate the effect size corresponding to known clinical efficacy.

Bibliography

1. Hay, M., Thomas, D. W., Craighead, J. L., Economides, C. & Rosenthal, J. Clinical development success rates for investigational drugs. *Nature Biotechnology* **32,** 40–51. ISSN: 1087-0156, 1546-1696. http://www.nature.com/articles/nbt.2786 (2020) (Jan. 2014).
2. Smietana, K., Siatkowski, M. & Møller, M. Trends in Clinical Success Rates. *Nature Reviews Drug Discovery* **15,** 379–380. ISSN: 1474-1776, 1474-1784. (2025) (June 2016).
3. Schuhmacher, A., Hinder, M., Brief, E., Gassmann, O. & Hartl, D. Benchmarking R&D Success Rates of Leading Pharmaceutical Companies: An Empirical Analysis of FDA Approvals (2006–2022). *Drug Discovery Today* **30,** 104291. ISSN: 13596446. (2025) (Feb. 2025).
4. Wouters, O. J., McKee, M. & Luyten, J. Estimated Research and Development Investment Needed to Bring a New Medicine to Market, 2009-2018. *JAMA* **323,** 844. ISSN: 0098-7484. https://jamanetwork.com/journals/jama/fullarticle/2762311 (2020) (Mar. 3, 2020).
5. Cook, D. *et al.* Lessons learned from the fate of AstraZeneca's drug pipeline: a five-dimensional framework. *Nature Reviews Drug Discovery* **13.** Bandiera_abtest: a Cg_type: Nature Research Journals Number: 6 Primary_atype: Reviews Publisher: Nature Publishing Group Subject_term: Business strategy in drug development Subject_term_id: business-strategy-in-drug-development, 419–431. ISSN: 1474-1784. https://www.nature.com/articles/nrd4309 (2021) (June 2014).
6. Of the Commissioner, O. Step 3: Clinical Research. *FDA.* (2025) (Tue, 11/18/2025 - 07:45).
7. Sertkaya, A., Beleche, T., Jessup, A. & Sommers, B. D. Costs of Drug Development and Research and Development Intensity in the US, 2000-2018. *JAMA Network Open* **7,** e2415445. ISSN: 2574-3805. (2025) (June 2024).
8. Scannell, J. W., Blanckley, A., Boldon, H. & Warrington, B. Diagnosing the Decline in Pharmaceutical R&D Efficiency. *Nature Reviews Drug Discovery* **11,** 191–200. ISSN: 1474-1776, 1474-1784. (2025) (Mar. 2012).
9. Bowling, H., Cocucci, A., Koo, D. C. E. & Harrison, R. K. Analysis of Phase II and Phase III Clinical Trial Terminations from 2013 to 2023. *Nature Reviews Drug Discovery.* ISSN: 1474-1784. (2025) (Dec. 2025).
10. Wu, S. S. *et al.* Reviving an R&D pipeline: a step change in the Phase II success rate. *Drug Discovery Today* **26,** 308–314. ISSN: 1359-6446. https://www.sciencedirect.com/science/article/pii/S1359644620304360 (2021) (Feb. 1, 2021).

11. Cooper, M. M., Loiseau, C., McCarthy, J. S. & Doolan, D. L. Human challenge models: tools to accelerate the development of malaria vaccines. *Expert Review of Vaccines* **18,** 241–251. ISSN: 1744-8395 (Mar. 2019).

12. Lambkin-Williams, R., Noulin, N., Mann, A., Catchpole, A. & Gilbert, A. S. The human viral challenge model: accelerating the evaluation of respiratory antivirals, vaccines and novel diagnostics. *Respiratory Research* **19,** 123. ISSN: 1465-9921. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6013893/ (2021) (2018).

13. McGonigle, P. & Ruggeri, B. Animal models of human disease: challenges in enabling translation. *Biochemical Pharmacology* **87,** 162–171. ISSN: 1873-2968 (Jan. 1, 2014).

14. Akhtar, A. The flaws and human harms of animal experimentation. *Cambridge quarterly of healthcare ethics: CQ: the international journal of healthcare ethics committees* **24,** 407–419. ISSN: 1469-2147 (Oct. 2015).

15. Begley, C. G. & Ellis, L. M. Drug development: Raise standards for preclinical cancer research. *Nature* **483,** 531–533. ISSN: 1476-4687 (Mar. 28, 2012).

16. Swaminathan, A., Lucas, R. M., Dear, K. & McMichael, A. J. Keyhole limpet haemocyanin - a model antigen for human immunotoxicological studies. *British Journal of Clinical Pharmacology* **78,** 1135–1142. ISSN: 1365-2125 (Nov. 2014).

17. Drennan, P. G. *et al.* Immunogenicity of MVA-BN Vaccine Deployed as Mpox Prophylaxis: A Prospective, Single-Centre, Cohort Study and Analysis of Transcriptomic Predictors of Response. *The Lancet Microbe* **6,** 101045. ISSN: 26665247. (2025) (June 2025).

18. UKHSA. Recommendations for the use of pre and post exposure vaccination during a monkeypox incident, 33. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1083791/Recommendations-for-pre-and-post-exposure-vaccination-during-a-monkeypox-incident-17-june-2022.pdf (2022) (June 17, 2022).

19. Otter, A. D. *et al.* Monkeypox Virus-Infected Individuals Mount Comparable Humoral Immune Responses as Smallpox-vaccinated Individuals. *Nature Communications* **14,** 5948. ISSN: 2041-1723. (2023) (Sept. 2023).

20. Jones, S. *et al. MpoxPlex: A High-Throughput and Versatile Multiplexed Immunoassay for Assessing and Discriminating between IgG Responses to Mpox Infection and Vaccination* June 2024. (2024).

21. Grifoni, A. *et al.* Defining Antigen Targets to Dissect Vaccinia Virus and Monkeypox Virus-Specific T Cell Responses in Humans. *Cell Host & Microbe* **30,** 1662–1670.e4. ISSN: 1931-3128. (2023) (Dec. 2022).

22. Satti, I. *et al.* Safety and Immunogenicity of a Candidate Tuberculosis Vaccine MVA85A Delivered by Aerosol in BCG-vaccinated Healthy Adults: A Phase 1, Double-Blind, Randomised Controlled Trial. *The Lancet Infectious Diseases* **14,** 939–946 (Oct. 2014).

23. da Silva Antunes, R. *et al.* The MegaPool Approach to Characterize Adaptive CD4+ and CD8+ T Cell Responses. *Current Protocols* **3,** e934. ISSN: 2691-1299 (Nov. 2023).

24. Vita, R. *et al.* The Immune Epitope Database (IEDB): 2018 Update. *Nucleic Acids Research* **47,** D339–D343. ISSN: 1362-4962 (Jan. 2019).

25. Minhaj, F. S. *et al.* Monkeypox Outbreak - Nine States, May 2022. *MMWR. Morbidity and mortality weekly report* **71,** 764–769. ISSN: 1545-861X (June 2022).

26. Rowland, R. *et al.* Safety and Immunogenicity of an FP9-vectored Candidate Tuberculosis Vaccine (FP85A), Alone and with Candidate Vaccine MVA85A in BCG-vaccinated Healthy Adults: A Phase I Clinical Trial. *Human vaccines & Immunotherapeutics* **9,** 50–62 (Jan. 2013).

27. Moraes-Cardoso, I. *et al.* Immune Responses Associated with Mpox Viral Clearance in Men with and without HIV in Spain: A Multisite, Observational, Prospective Cohort Study. *The Lancet Microbe* **5.** ISSN: 2666-5247. (2024) (Aug. 2024).

28. Lemieux, A. *et al.* Enhanced Detection of Antigen-Specific T Cells by a Multiplexed AIM Assay. *Cell Reports Methods* **4,** 100690. ISSN: 2667-2375 (Jan. 2024).

29. Brodin, P. *et al.* Variation in the human immune system is largely driven by non-heritable influences. *Cell* **160,** 37–47. ISSN: 1097-4172 (Jan. 15, 2015).

30. Yang, N. *et al.* Heat-Inactivated Modified Vaccinia Virus Ankara Boosts Th1 Cellular and Humoral Immunity as a Vaccine Adjuvant. *NPJ vaccines* **7,** 120. ISSN: 2059-0105 (Oct. 2022).

31. Wang, W. *et al.* Elucidating Mechanisms of Antitumor Immunity Mediated by Live Oncolytic Vaccinia and Heat-Inactivated Vaccinia. *Journal for ImmunoTherapy of Cancer* **9.** ISSN: 2051-1426. (2026) (Sept. 2021).

32. Skeldon, A. M. *et al.* Structural Insight into the cGAS Active Site Explains Differences between Therapeutically Relevant Species. *Communications Chemistry* **8,** 88. ISSN: 2399-3669. (2026) (Mar. 2025).

33. Registry, I. *A Phase I Randomized Study of IMSB301 in Healthy Volunteers* https://www.isrctn.com (2026).

34. Decout, A., Katz, J. D., Venkatraman, S. & Ablasser, A. The cGAS–STING Pathway as a Therapeutic Target in Inflammatory Diseases. *Nature Reviews Immunology* **21,** 548–569. ISSN: 1474-1741. (2026) (Sept. 2021).

35. Gao, J. *et al.* CDK Inhibitor Palbociclib Targets STING to Alleviate Autoinflammation. *EMBO reports* **23,** e53932. ISSN: 1469-3178. PMID: 35403787 (June 7, 2022).

36. Cho, Y. M. & Furie, R. The Development of Litifilimab (BIIB 059) for Cutaneous and Systemic Lupus Erythematosus. *Immunotherapy* **16,** 15–20. ISSN: 1750-7448 (Jan. 2024).

37. Amgen. *A Phase 2, Multicenter, Randomized, Double-Blind, Placebo-Controlled, Parallel-Group Trial to Investigate the Efficacy and Safety of Daxdilimab Subcutaneous Injection in Reducing Disease Activity in Adult Participants With Moderate-to-Severe Primary Discoid Lupus Erythematosus* Clinical Trial Registration NCT05591222 (clinicaltrials.gov, Oct. 2025). (2026).

38. Drennan, P. G., Karponis, D., Richards, D., Coles, M. & Fullerton, J. N. In Vivo Human Keyhole Limpet Hemocyanin Challenge in Early Phase Drug Development: A Systematic Review. *Clinical and Translational Science* **Mar;16(3):357-382.** ISSN: 1752-8062. (2023).

39. Oyelaran, O. & Gildersleeve, J. C. Evaluation of human antibody responses to keyhole limpet hemocyanin on a carbohydrate microarray. *Proteomics Clinical Applications* **4,** 285–294 (Mar. 2010).

40. Harris, J. R. & Markl, J. Keyhole limpet hemocyanin (KLH): a biomedical review. *Micron (Oxford, England: 1993)* **30,** 597–623. ISSN: 0968-4328 (Dec. 1999).

41. Salvaggio, J., Castro-Murillo, E. & Kundur, V. Immunologic response of atopic and normal individuals to keyhole limpet hemocyanin. *Journal of Allergy* **44,** 344–354. ISSN: 0021-8707 (Dec. 1969).

42. Lebrec, H. *et al.* An inter-laboratory retrospective analysis of immunotoxicological endpoints in non-human primates: T-cell-dependent antibody responses. *Journal of Immunotoxicology* **8,** 238–250. ISSN: 1547-6901 (Sept. 2011).

43. Lebrec, H. *et al.* T-cell-dependent antibody responses in the rat: Forms and sources of keyhole limpet hemocyanin matter. *Journal of Immunotoxicology* **11.** Publisher: Taylor & Francis _eprint: https://doi.org/10.3109/1547691X.2013.822948, 213–221. ISSN: 1547-691X. https://doi.org/10.3109/1547691X.2013.822948 (2021) (July 1, 2014).

44. Kawai, R. *et al.* Evaluation of canine T-cell dependent antibody response to the primary and secondary immunization with keyhole limpet hemocyanin. *The Journal of Toxicological Sciences* **38,** 571–579. ISSN: 1880-3989 (2013).

45. Curtis, J. E., Hersh, E. M., Butler, W. T. & Rossen, R. D. Antigen dose in the human immune response: Dose-response relationships in the human immune response to Keyhole limpet hemocyanin. *The Journal of Laboratory and Clinical Medicine* **78.** Publisher: Elsevier, 61–69. ISSN: 0022-2143, 1532-6543. https://www.translationalres.com/article/0022-2143(71)90204-6/abstract (2020) (July 1, 1971).

46. Curtis, J. E., Hersh, E. M., Harris, J. E., McBride, C. & Freireich, E. J. The human primary immune response to keyhole limpet haemocyanin: interrelationships of delayed hypersensitivity, antibody response and in vitro blast transformation. *Clinical and Experimental Immunology* **6,** 473–491. ISSN: 0009-9104 (Apr. 1970).

47. Korver, K., Boeschoten, E. W., Krediet, R. T., van Steenis, G. & Schellekens, P. T. Dose-response effects in immunizations with keyhole limpet haemocyanin and rabies vaccine: shift in some immunodeficiency states. *Clinical and Experimental Immunology* **70,** 328–335. ISSN: 0009-9104. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1542077/ (2020) (Nov. 1987).

48. Korver, K., Zeijlemaker, W. P., Schellekens, P. T. A. & Vossen, J. M. Measurement of primary in vivo IgM- and IgG-antibody response to KLH in humans: Implications of pre-immune IgM binding in antigen-specific ELISA. *Journal of Immunological Methods* **74,** 241–251. ISSN: 0022-1759. http://www.sciencedirect.com/science/article/pii/0022175984902916 (2020) (Nov. 30, 1984).

49. Page, M. J. *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* **372.** Publisher: British Medical Journal Publishing Group Section: Research Methods & Reporting, n71. ISSN: 1756-1833. https://www.bmj.com/content/372/bmj.n71 (2022) (Mar. 29, 2021).

50. Ouzzani, M., Hammady, H., Fedorowicz, Z. & Elmagarmid, A. Rayyan—a web and mobile app for systematic reviews. *Systematic Reviews* **5,** 210. ISSN: 2046-4053. https://doi.org/10.1186/s13643-016-0384-4 (2022) (Dec. 5, 2016).

51. Hoffmann, T. C. *et al.* Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* **348.** Publisher: British Medical Journal Publishing Group Section: Research Methods & Reporting, g1687. ISSN: 1756-1833. https://www.bmj.com/content/348/bmj.g1687 (2022) (Mar. 7, 2014).

52. Abrams, J. R. *et al.* CTLA4Ig-mediated blockade of T-cell costimulation in patients with psoriasis vulgaris. *Journal of Clinical Investigation* **103,** 1243–1252. ISSN: 0021-9738. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC408469/ (2020) (May 1, 1999).

53. Van der Kolk, L. E., Baars, J. W., Prins, M. H. & van Oers, M. H. J. Rituximab treatment results in impaired secondary humoral immune responsiveness. *Blood* **100.** Institution: van der Kolk, Lizet E. Department of Hematology, Academic Medical Center, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands., 2257–2259 (Sept. 2002).

54. Bingham, C. O. 3. *et al.* Immunization responses in rheumatoid arthritis patients treated with rituximab: results from a controlled clinical trial. *Arthritis & Rheumatism* **62,** 64–74 (Jan. 2010).

55. Smith, K. *et al.* Treatment with Recombinant Growth Hormone Is Associated with Modest Improvement in CD4 Lymphocyte Reconstitution in HIV-Infected Persons on Antiretroviral Therapy: Results of ACTG A5174. *AIDS Research and Human Retroviruses* **26,** 425–432. ISSN: 0889-2229. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2864047/ (2022) (Apr. 2010).

56. Curti, B. D. *et al.* OX40 is a potent immune stimulating target in late stage cancer patients. *Cancer research* **73,** 7189–7198. ISSN: 0008-5472. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3922072/ (2022) (Dec. 15, 2013).

57. Poirier, N. *et al.* First-in-Human Study in Healthy Subjects with FR104, a Pegylated Monoclonal Antibody Fragment Antagonist of CD28. *Journal of Immunology* **197.** Institution: Vanhove, Bernard. LabEx ImmunoGraft Oncology, Nantes F44000, Nantes, France; and., 4593–4602 (2016).

58. Shi, R. *et al.* Pharmacokinetic, Pharmacodynamic, and Safety Profile of a Novel Anti-CD28 Domain Antibody Antagonist in Healthy Subjects. *Journal of Clinical Pharmacology* **57,** 161–172. ISSN: 0091-2700. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5697635/ (2021) (Feb. 2017).

59. Sullivan, B. A. *et al.* Inducible T-cell co-stimulator ligand (ICOSL) blockade leads to selective inhibition of anti-KLH IgG responses in subjects with systemic lupus erythematosus. *Lupus Science & Medicine* **3.** ISSN: 2053-8790. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4836284/ (2020) (Apr. 8, 2016).

60. St Clair, E. W. *et al.* The Clinical Efficacy and Safety of Baminercept, a Lymphotoxin-Receptor Fusion Protein, in Primary Sjögren's Syndrome: Results from a Randomized, Double-Blind, Placebo-Controlled Phase II Trial. *Arthritis & rheumatology (Hoboken, N.J.)* **70,** 1470–1480. ISSN: 2326-5191. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6115299/ (2022) (Sept. 2018).

61. Karnell, J. L. *et al.* A CD40L-targeting protein reduces autoantibodies and improves disease activity in patients with autoimmunity. *Science Translational Medicine* **11** (2019).

62. Espié, P. *et al.* First-in-human clinical trial to assess pharmacokinetics, pharmacodynamics, safety, and tolerability of iscalimab, an anti-CD40 monoclonal antibody. *American Journal of Transplantation: Official Journal of the American Society of Transplantation and the American Society of Transplant Surgeons* **20,** 463–473. ISSN: 1600-6143 (2020).

63. Yang, J. *et al.* First-in-human study of the safety, tolerability, pharmacokinetics, and pharmacodynamics of ALPN-101, a dual CD28/ICOS antagonist, in healthy adult subjects. *Clinical and translational science* **14.** Institution: Peng, Stanford L. Alpine Immune Sciences, Inc., Seattle, Washington, USA., 1314–1326 (2021).

64. Saghari, M. *et al.* OX40L Inhibition Suppresses KLH-driven Immune Responses in Healthy Volunteers: A Randomized Controlled Trial Demonstrating Proof-of-Pharmacology for KY1005. *Clinical Pharmacology and Therapeutics.* ISSN: 1532-6535 (Jan. 29, 2022).

65. Rentenaar, R. J. *et al.* Immune responsiveness in renal transplant recipients: mycophenolic acid severely depresses humoral immunity in vivo. *Kidney International* **62,** 319–328. ISSN: 0085-2538 (July 2002).

66. Saville, M. W. *et al.* Evaluation of Humoral Immune Responses to Vaccination with Tetanus Toxoid and KLH in Rituximab-Treated Follicular Non-Hodgkin's Lymphoma Patients Compared to Healthy Volunteers. *Blood* **112,** 3066. ISSN: 0006-4971. https://doi.org/10.1182/blood.V112.11.3066.3066 (2022) (Nov. 16, 2008).

67. Struijk, G. H. *et al.* Maintenance immunosuppressive therapy with everolimus preserves humoral immune responses. *Kidney International* **78,** 934–940 (Nov. 2010).

68. Boulton, C., Meiser, K., David, O. J. & Schmouder, R. Pharmacodynamic Effects of Steady-State Fingolimod on Antibody Response in Healthy Volunteers: A 4-Week, Randomized, Placebo-Controlled, Parallel-Group, Multiple-Dose Study. *The Journal of Clinical Pharmacology* **52.** _eprint: https://accp1.onlinelibrary.wiley.com/doi/pdf/10.1177/00912 1879–1890. ISSN: 1552-4604. https://accp1.onlinelibrary.wiley.com/doi/abs/10.1177/0091270011427908 (2020) (2012).

69. Kaufman, M. *et al.* Natalizumab treatment shows no clinically meaningful effects on immunization responses in patients with relapsing-remitting multiple sclerosis. *Journal of the Neurological Sciences* **341,** 22–27 (June 2014).

70. Bar-Or, A. *et al.* Effect of ocrelizumab on vaccine responses in patients with multiple sclerosis: The VELOCE study. *Neurology* **95.** Institution: Traboulsee, Anthony. From the Department of Neurology and Center for Neuroinflammation and Experimental Therapeutics (A.B.-O.), Perelman School of Medicine, University of Pennsylvania, Philadelphia; The Minneapolis Clinic of Neurology (J.C.C.), MN; F. Hoffmann-La Roche Ltd (C.C., J.E., M.M., D.S.), Basel, Switzerland; Central Texas Neurology Consultants (E.J.F.), Round Rock; Genentech, Inc (A.H.), South San Francisco, CA; John McNamara Consulting Ltd (J.M.), Cambridge, UK; Department of Neurology (D.S.R.), Multiple Sclerosis Division, University of South Florida College of Medicine, Tampa; Territory Neurology and Research Institution (J.K.W.), Tucson, AZ; Division of Infectious Diseases (K.L.W.), Oregon Health & Science University, Portland; and University of British Columbia (A.T.), Vancouver, Canada. (2020).

71. Kondratenko, I., Amlot, P. L., Webster, A. D. & Farrant, J. Lack of specific antibody response in common variable immunodeficiency (CVID) associated with failure in production of antigen-specific memory T cells. MRC Immunodeficiency Group. *Clinical and Experimental Immunology* **108,** 9–13. ISSN: 0009-9104 (Apr. 1997).

72. Kuijpers, T. W., Etzioni, A., Pollack, S. & Pals, S. T. Antigen-specific immune responsiveness and lymphocyte recruitment in leukocyte adhesion deficiency type II. *International Immunology* **9.** Institution: Kuijpers, T W. Department of Pediatrics, University of Amsterdam, The Netherlands., 607–613. ISSN: 0953-8178 (Apr. 1997).

73. Suchin, K. R. *et al.* Extracorporeal photochemotherapy does not suppress T- or B-cell responses to novel or recall antigens. *Journal of the American Academy of Dermatology* **41,** 980–986 (Dec. 1999).

74. Valdez, H. *et al.* Response to immunization with recall and neoantigens after prolonged administration of an HIV-1 protease inhibitor-containing regimen. ACTG 375 team. AIDS Clinical Trials Group. *AIDS (London, England)* **14,** 11–21. ISSN: 0269-9370 (Jan. 7, 2000).

75. Markert, M. L. *et al.* Effect of highly active antiretroviral therapy and thymic transplantation on immunoreconstitution in HIV infection. *AIDS Research & Human Retroviruses* **16.** Institution: Markert, M L. Department of Pediatrics, Duke University Medical Center, Durham, North Carolina 27710, USA. marke001@mc.duke.edu, 403–413. ISSN: 0889-2229 (Mar. 2000).

76. Lange, C. G. *et al.* Nadir CD4+ T-cell count and numbers of CD28+ CD4+ T-cells predict functional responses to immunizations in chronic HIV-1 infection. *AIDS* **17.** Institution: Lange, Christoph G. Center for AIDS Research, Case Western University School of Medicine, Cleveland, OH 44106, USA., 2015–2023 (Sept. 2003).

77. Lange, C. G. *et al.* Proliferation responses to HIVp24 during antiretroviral therapy do not reflect improved immune phenotype or function. *AIDS* **18,** 605–613 (Mar. 2004).

78. Boelens, P. G. *et al.* Glutamine-enriched enteral nutrition increases in vitro interferon-gamma production but does not influence the in vivo specific antibody response to KLH after severe trauma. A prospective, double blind, randomized clinical study. *Clinical Nutrition (Edinburgh, Scotland)* **23,** 391–400. ISSN: 0261-5614 (June 2004).

79. Smith, T. P., Kennedy, S. L. & Fleshner, M. Influence of age and physical activity on the primary in vivo antibody and T cell-mediated responses in men. *Journal of Applied Physiology* **97,** 491–498 (Aug. 2004).

80. Boelens, P. G. *et al.* Primary immune response to keyhole limpet haemocyanin following trauma in relation to low plasma glutamine. *Clinical and Experimental Immunology* **136,** 356–364. ISSN: 0009-9104 (May 2004).

81. Smith, A. *et al.* The relationship between distress and the development of a primary immune response to a novel antigen. *Brain, Behavior, & Immunity* **18,** 65–75 (Jan. 2004).

82. Smith, A. J., Vollmer-Conna, U., Bennett, B., Hickie, I. B. & Lloyd, A. R. Influences of distress and alcohol consumption on the development of a delayed-type hypersensitivity skin test response. *Psychosomatic Medicine* **66,** 614–619 (Aug. 2004).

83. Miller, J. S. *et al.* Diminished neo-antigen response to keyhole limpet hemocyanin (KLH) vaccines in patients after treatment with chemotherapy or hematopoietic cell transplantation. *Clinical Immunology (Orlando, Fla.)* **117,** 144–151. ISSN: 1521-6616 (Nov. 2005).

84. Grant, R. *et al.* Cardiovascular Exercise Intervention Improves the Primary Antibody Response to Keyhole Limpet Hemocyanin (KLH) in Previously Sedentary Older Adults. *Brain, behavior, and immunity* **22,** 923–932. ISSN: 0889-1591. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2576741/ (2020) (Aug. 2008).

85. Spazierer, D. *et al.* T helper 2 biased de novo immune response to Keyhole Limpet Hemocyanin in humans. *Clinical & Experimental Allergy* **39.** Institution: Spazierer, D. Department of Dermatology, Division of Immunology, Allergy and Infectious Diseases, Medical University of Vienna, Vienna Competence Center, 1090 Vienna, Austria., 999–1008 (July 2009).

86. Rodriguez, B. *et al.* In vitro naive T cell proliferation failure predicts poor post-immunization responses to neoantigen, but not recall antigens, in HIV-infection. *Clinical Immunology* **136,** 400–408 (Sept. 2010).

87. Milgrom, H. *et al.* Response to Cutaneous Immunization with Low-Molecular-Weight Subunit Keyhole Limpet Hemocyanin. *International Archives of Allergy and Immunology* **157,** 269–274. ISSN: 1018-2438. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3224512/ (2020) (Feb. 2012).

88. Ferbas, J. *et al.* A novel assay to measure B cell responses to keyhole limpet haemocyanin vaccination in healthy volunteers and subjects with systemic lupus erythematosus. *British Journal of Clinical Pharmacology* **76,** 188–202. ISSN: 1365-2125 (Aug. 2013).

89. Gallegos, A. M. *et al.* Toward identifying the effects of the specific components of Mindfulness-Based Stress Reduction on biologic and emotional outcomes among older adults. *Journal of Alternative* **19,** 787–792 (Oct. 2013).

90. Moynihan, J. A. *et al.* Mindfulness-based stress reduction for older adults: effects on executive function, frontal alpha asymmetry and immune function. *Neuropsychobiology* **68,** 34–43 (2013).

91. Belson, A. *et al.* Characterisation of the clinical and activated T cell response to repeat delayed-type hypersensitivity skin challenges in human subjects, with KLH and PPD, as a potential model to test T cell-targeted therapies. *Inflammation Research* **65,** 389–404 (May 2016).

92. Giesecke, C. *et al.* Simultaneous Presence of Non- and Highly Mutated Keyhole Limpet Hemocyanin (KLH)-Specific Plasmablasts Early after Primary KLH Immunization Suggests Cross-Reactive Memory B Cell Activation. *Journal of Immunology (Baltimore, Md.: 1950)* **200,** 3981–3992. ISSN: 1550-6606 (2018).

93. Swaminathan, A. *et al.* Exposure to Solar UVR Suppresses Cell-Mediated Immunization Responses in Humans: The Australian Ultraviolet Radiation and Immunity Study. *Journal of Investigative Dermatology* **139.** Institution: Lucas, Robyn M. National Centre for Epidemiology and Population Health, Research School of Population Health, The Australian National University, Canberra, Australia; Centre for Ophthalmology and Visual Science, University of Western Australia, Perth, Australia. Electronic address: robyn.lucas@anu.edu.au., 1545–1553 (2019).

94. Saghari, M. *et al.* A randomized controlled trial with a delayed-type hypersensitivity model using keyhole limpet haemocyanin to evaluate adaptive immune responses in man. *British Journal of Clinical Pharmacology,* bcp.14588. ISSN: 0306-5251, 1365-2125. https://onlinelibrary.wiley.com/doi/10.1111/bcp.14588 (2020) (Oct. 28, 2020).

95. Otterhaug, T. *et al.* Photochemical Internalization Enhanced Vaccination Is Safe, and Gives Promising Cellular Immune Responses to an HPV Peptide-Based Vaccine in a Phase I Clinical Study in Healthy Volunteers. *Frontiers in Immunology* **11,** 576756. ISSN: 1664-3224 (2020).

96. Dickson, M. *et al.* Immunisation challenges with keyhole limpet haemocyanin (KLH) and bacteriophage PhiX174: potential for modelling in vivo pharmacodynamic effects. *Immunology* **143,** 70–70 (Dec. 1, 2014).

97. Jain, A. *et al.* Partial immune reconstitution of X-linked hyper IgM syndrome with recombinant CD40 ligand. *Blood* **118,** 3811–3817. ISSN: 0006-4971. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3193261/ (2022) (Oct. 6, 2011).

98. Saghari, M. *et al.* OX40L blocking monoclonal antibody KY1005 strongly suppresses the delayedtype hypersensitivity skin response to klh in healthy volunteers. *Journal of the dermatology nurses' association* **12.** https://www.cochranelibrary.com/central/doi/10.1002/central/CN-02278646/full (2020).

99. Sokal, J. E. Editorial: Measurement of delayed skin-test responses. *The New England Journal of Medicine* **293,** 501–502. ISSN: 0028-4793 (Sept. 4, 1975).

100. Yang, J. *et al.* A double blind, placebo controlled, single ascending dose (SAD) and multiple ascending dose (MAD) study of ALPN-101, a first-inclass dual ICOS/CD28 antagonist, in healthy volunteers (HV). *Annals of the rheumatic diseases* **79,** 1040–. https://www.cochranelibrary.com/central/doi/10.1002/central/CN-02236034/full (SUPPL 1 2020).

101. Shi, R. *et al.* A novel anti-CD28 domain antibody antagonist shows a favorable pharmacokinetic, pharmacodynamic and safety profile. *Arthritis & rheumatology* **67.** https://www.cochranelibrary.com/central/doi/10.1002/central/CN-01198136/full (no pagination 2015).

102. Suchard, S. J. *et al.* A monovalent anti-human CD28 domain antibody antagonist: preclinical efficacy and safety. *Journal of Immunology (Baltimore, Md.: 1950)* **191,** 4599–4610. ISSN: 1550-6606 (Nov. 1, 2013).

103. Kraus, T. A., Cheifetz, A., Toy, L., Meddings, J. B. & Mayer, L. Evidence for a genetic defect in oral tolerance induction in inflammatory bowel disease. *Inflammatory Bowel Diseases* **12,** 82–88 (Feb. 2006).

104. Chan, A.-W. & Altman, D. G. Epidemiology and reporting of randomised trials published in PubMed journals. *Lancet (London, England)* **365,** 1159–1162. ISSN: 1474-547X (Apr. 26, 2005).

105. Von Elm, E. *et al.* Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ : British Medical Journal* **335,** 806–808. ISSN: 0959-8138. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2034723/ (2022) (Oct. 20, 2007).

106. Chan, A.-W. *et al.* SPIRIT 2013 Statement: Defining Standard Protocol Items for Clinical Trials. *Annals of internal medicine* **158,** 200–207. ISSN: 0003-4819. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5114123/ (2022) (Feb. 5, 2013).

107. Moher, D., Schulz, K. F. & Altman, D. G. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet (London, England)* **357,** 1191–1194. ISSN: 0140-6736 (Apr. 14, 2001).

108. Yap, C. *et al.* The need for reporting guidelines for early phase dose-finding trials: Dose-Finding CONSORT Extension. *Nature Medicine* **28.** Number: 1 Publisher: Nature Publishing Group, 6–7. ISSN: 1546-170X. https://www.nature.com/articles/s41591-021-01594-1 (2022) (Jan. 2022).

109. De Maeyer, R. P. H. *et al.* Blocking elevated p38 MAPK restores efferocytosis and inflammatory resolution in the elderly. *Nature Immunology* **21,** 615–625. ISSN: 1529-2916 (June 2020).

110. Akbar, A. N. *et al.* Investigation of the cutaneous response to recall antigen in humans in vivo. *Clinical and Experimental Immunology* **173,** 163–172. ISSN: 1365-2249 (Aug. 2013).

111. Del Giudice, G., Rappuoli, R. & Didierlaurent, A. M. Correlates of adjuvanticity: A review on adjuvants in licensed vaccines. *Seminars in Immunology* **39,** 14–21. ISSN: 1096-3618 (Oct. 2018).

112. Bouros, D., Zeros, G., Panaretos, C., Vassilatos, C. & Siafakas, N. Palpation vs Pen Method for the Measurement of Skin Tuberculin Reaction (Mantoux Test). *Chest* **99,** 416–419. ISSN: 0012-3692 (Feb. 1991).

113. Mangiola, S. *et al.* Sccomp: Robust Differential Composition and Variability Analysis for Single-Cell Data. *Proceedings of the National Academy of Sciences* **120,** e2203828120. ISSN: 0027-8424, 1091-6490. (2025) (Aug. 2023).

114. Moysi, E., Paris, R. M., Le Grand, R., Koup, R. A. & Petrovas, C. Human lymph node immune dynamics as driver of vaccine efficacy: an understudied aspect of immune responses. *Expert Review of Vaccines* **21.** Publisher: Taylor & Francis, 633–644. ISSN: 1476-0584. https://ezproxy-prd.bodleian.ox.ac.uk:2398/doi/full/10.1080/14760584.2022.2045198 (2022) (May 4, 2022).

115. Havenar-Daughton, C. *et al.* Direct probing of germinal center responses reveals immunological features and bottlenecks for neutralizing antibody responses to HIV Env trimer. *Cell reports* **17,** 2195–2209. ISSN: 2211-1247. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5142765/ (2021) (Nov. 22, 2016).

116. Havenar-Daughton, C. *et al.* Normal human lymph node T follicular helper cells and germinal center B cells accessed via fine needle aspirations. *Journal of Immunological Methods* **479,** 112746. ISSN: 1872-7905 (Apr. 2020).

117. Turner, J. S. *et al.* SARS-CoV-2 mRNA vaccines induce persistent human germinal centre responses. *Nature* **596.** Bandiera_abtest: a Cg_type: Nature Research Journals Number: 7870 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: B cells;Germinal centres;RNA vaccines;SARS-CoV-2 Subject_term_id: b-cells;germinal-centres;rna-vaccines;sars-cov-2, 109–113. ISSN: 1476-4687. https://www.nature.com/articles/s41586-021-03738-2 (2021) (Aug. 2021).

118. Win, Z. *et al.* Systematic Evaluation of Kinetics and Distribution of Muscle and Lymph Node Activation Measured by 18F-FDG- and 11C-PBR28-PET/CT Imaging, and Whole Blood and Muscle Transcriptomics After Immunization of Healthy Humans With Adjuvanted and Unadjuvanted Vaccines. *Frontiers in Immunology* **11,** 3863. ISSN: 1664-3224. https://www.frontiersin.org/article/10.3389/fimmu.2020.613496 (2021) (2021).

119. Stride, E. & Saffari, N. Microbubble ultrasound contrast agents: a review. *Proceedings of the Institution of Mechanical Engineers. Part H, Journal of Engineering in Medicine* **217,** 429–447. ISSN: 0954-4119 (2003).

120. Blomley, M. J. K., Cooke, J. C., Unger, E. C., Monaghan, M. J. & Cosgrove, D. O. Microbubble contrast agents: a new era in ultrasound. *BMJ : British Medical Journal* **322,** 1222–1225. ISSN: 0959-8138. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1120332/ (2022) (May 19, 2001).

121. Goldberg, B. B., Merton, D. A., Liu, J.-B., Murphy, G. & Forsberg, F. Contrast-Enhanced Sonographic Imaging of Lymphatic Channels and Sentinel Lymph Nodes. *Journal of Ultrasound in Medicine* **24.** _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.7863/jum 953–965. ISSN: 1550-9613. https://onlinelibrary.wiley.com/doi/abs/10.7863/jum.2005.24.7.953 (2022) (2005).

122. Wei, Y. *et al.* Combination of Lymphatic and Intravenous Contrast-Enhanced Ultrasound for Evaluation of Cervical Lymph Node Metastasis from Papillary Thyroid Carcinoma: A Preliminary Study. *Ultrasound in Medicine & Biology* **47,** 252–260. ISSN: 1879-291X (Feb. 2021).

123. Goldberg, B. B. *et al.* Sentinel lymph nodes in a swine model with melanoma: contrast-enhanced lymphatic US. *Radiology* **230,** 727–734. ISSN: 0033-8419 (Mar. 2004).

124. Gkegkes, I. D. & lavazzo, C. Contrast Enhanced Ultrasound (CEU) Using Microbubbles for Sentinel Lymph Node Biopsy in Breast Cancer: a Systematic Review. *Acta Chirurgica Belgica* **115,** 212–218. ISSN: 0001-5458. http://www.tandfonline.com/doi/full/10.1080/00015458.2015.11681099 (2022) (Jan. 2015).

125. Machado, P. *et al.* Subdermal Ultrasound Contrast Agent Injection for Sentinel Lymph Node Identification: An Analysis of Safety and Contrast Agent Dose in Healthy Volunteers. *Journal of Ultrasound in Medicine* **37,** 1611–1620. ISSN: 0278-4297, 1550-9613. https://onlinelibrary.wiley.com/doi/10.1002/jum.14502 (2022) (July 2018).

126. Menshykau, D. & Tanaka, S. Mechanistic Image-Based Modelling: Concepts and Applications. [Review]. *Handbook of Experimental Pharmacology* **1,** 231–261. ISSN: 0171-2004 (2019).

127. Nielsen Moody, A., Cox, K., Haigh, I., Chen, Y. & Sharma, N. Does Contrast Enhanced Ultrasound (CEUS) of Normal/Benign Axillary Lymph Nodes in Patients with Breast Cancer Identify Significant Axillary Nodal Burden? *European Journal of Radiology* **132,** 109311. ISSN: 1872-7727 (Nov. 2020).

128. Sever, A. R. *et al.* Preoperative needle biopsy of sentinel lymph nodes using intradermal microbubbles and contrast-enhanced ultrasound in patients with breast cancer. *AJR. American journal of roentgenology* **199,** 465–470. ISSN: 1546-3141 (Aug. 2012).

129. Morgan, P. *et al.* Can the flow of medicines be improved? Fundamental pharmacokinetic and pharmacological principles toward improving Phase II survival. *Drug Discovery Today* **17,** 419–424. ISSN: 1359-6446. https://www.sciencedirect.com/science/article/pii/S1359644611004776 (2021) (May 1, 2012).

130. Brodin, P. & Davis, M. M. Human immune system variation. *Nature reviews. Immunology* **17,** 21–29. ISSN: 1474-1733. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5328245/ (2021) (Jan. 2017).

131. Lindley, D. V. *Making Decisions* 2. ed., repr. ISBN: 978-0-471-90803-6 978-0-471-90808-1 (Wiley, London, 1994).

132. Morris, T. P., White, I. R. & Crowther, M. J. Using simulation studies to evaluate statistical methods. *Statistics in Medicine* **38,** 2074–2102. ISSN: 1097-0258. https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.8086 (2019) (2019).

133. Gelman, A. *et al. Bayesian Data Analysis* Third edition. ISBN: 978-1-4398-4095-5 (CRC Press, Taylor & Francis Group, Boca Raton London New York, 2014).

134. Team, S. D. *Stan Modeling Language Users Guide and Reference Manual, Version 2.25* https://mc-stan.org/ (2020).

135. Zhang, L., Carpenter, B., Gelman, A. & Vehtari, A. Pathfinder: Parallel Quasi-Newton Variational Inference. *Journal of Machine Learning Research* **23,** 1–49. ISSN: 1533-7928. (2026) (2022).

136. Müller, P., Berry, D. A., Grieve, A. P. & Krams, M. A Bayesian Decision-Theoretic Dose-Finding Trial. *Decision Analysis* **3,** 197–207. ISSN: 1545-8490, 1545-8504. (2025) (Dec. 2006).

137. O'Hagan, A. & Stevens, J. W. A Framework for Cost-effectiveness Analysis from Clinical Trial Data. *Health Economics* **10,** 303–315. ISSN: 1057-9230, 1099-1050. (2025) (June 2001).

138. Gittins, J. C. & Pezeshk, H. A DECISION THEORETIC APPROACH TO SAMPLE SIZE DETERMINATION IN CLINICAL TRIALS. *Journal of Biopharmaceutical Statistics* **12,** 535–551. ISSN: 1054-3406, 1520-5711. (2025) (Jan. 2002).

139. Mellnik, J. & Scannell, J. *Tripartite Models for Estimating the Value of Drug Candidates and Decision Tools* Mar. 2025. arXiv: 2503.22117 [stat]. (2026).

140. Scannell, J. W. *et al.* Predictive Validity in Drug Discovery: What It Is, Why It Matters and How to Improve It. *Nature Reviews Drug Discovery* **21,** 915–931. ISSN: 1474-1784. (2026) (Dec. 2022).

141. Strong, M., Oakley, J. E., Brennan, A. & Breeze, P. Estimating the Expected Value of Sample Information Using the Probabilistic Sensitivity Analysis Sample. *Medical Decision Making* **35,** 570–583. ISSN: 0272-989X. (2025) (July 2015).

142. De Visser, S. J., Cohen, A. F. & Kenter, M. J. H. Integrating Scientific Considerations into R&D Project Valuation. *Nature Biotechnology* **38,** 14–18. ISSN: 1087-0156, 1546-1696. (2025) (Jan. 2020).

143. Harrison, M. & Lerer, L. Real Options for Biotechnology Valuation. *Nature Biotechnology* **20,** 223–223. ISSN: 1087-0156, 1546-1696. (2025) (Mar. 2002).

144. Williams, C. J., Wilson, K. J. & Wilson, N. A Comparison of Prior Elicitation Aggregation Using the Classical Method and SHELF. *Journal of the Royal Statistical Society Series A: Statistics in Society* **184,** 920–940. ISSN: 0964-1998, 1467-985X. (2025) (July 2021).

145. Alden, K. *et al.* Using argument notation to engineer biological simulations with increased confidence. *Journal of The Royal Society Interface* **12.** Publisher: Royal Society, 20141059. https://royalsocietypublishing.org/doi/10.1098/rsif.2014.1059 (2022) (Mar. 6, 2015).