Development and validation of instruments to assess public acceptance of alternatives to needle and syringe-based vaccination



Jonathan Kantor

St Edmund Hall

Supervised by:

Professor Robert C Carlisle

Dr Michael Morrison

Professor Samantha Vanderslott

Department of Engineering Science

University of Oxford

A thesis presented for the degree of Doctor of Philosophy

Abstract

Alternative vaccine delivery technologies such as transcutaneous ultrasound may hold promise to improve vaccination rates worldwide. Yet translational engineering efforts sometimes fail to yield marketable products due to alignment failure with end-user preferences, public sentiment, or stakeholder expectations and funding, and therefore a better understanding of public preferences for these technologies is critical. This thesis highlights a measurement-first approach to develop brief, psychometrically robust instruments to quantify human factors data early and iteratively, converting nebulous uncertainty regarding adoption likelihood into quantitative risk data.

This thesis outlines the development and validation of three instruments: the Oxford Needle Experience (ONE) scale, an instrument to assess public attitudes to needles; the Oxford Vaccine Hesitancy Scale (OVHS), a multidimensional scale to assess vaccine hesitancy; and the Oxford Benchmark Scale for Rating Vaccine Technologies (OBSRVT), a multidimensional instrument that compares responses to the benchmark of needle and syringe-based vaccination. These instruments, along with a range of measures including a discrete choice experiment, stated willingness to pay, and stated preferences, were then used to assess preferences for ultrasound-mediated vaccination and intranasal vaccination, exploring public enthusiasm for these next-generation technologies while also providing additional validity evidence for the underlying instruments.

Across demographically representative samples in the United Kingdom and United States (n=3,980), this thesis highlights the development and validation of these

instruments, which exhibited clean, multidimensional structures with strong reliability and validity evidence. In demographically representative populations in both the United Kingdom and United States, ultrasound-mediated vaccination is preferred over both intranasal and needle and syringe-based vaccination, and this preference is particularly pronounced among needle- and vaccine-hesitant individuals.

This thesis also outlines the development of a transferable approach to discrimination and cut-off determination applied to the OVHS, including the development of algorithms that could be broadly applicable to developing cut-off values when faced with large numbers of datasets—useful not only for instrument development, but for human and artificial intelligence decision-making under uncertainty more broadly. Ultimately, this thesis highlights preferences for ultrasound-mediated vaccination while also outlining a transferable approach for biotechnology de-risking.

Acknowledgments

This thesis truly would not have been possible without the help and dedication of a number of people, only some of whom are acknowledged here.

First, I would like to acknowledge and thank my supervisors, led by the brilliant, intrepid, and indefatigable Professor Bob Carlisle. Bob's enthusiasm and faith in me (and in this project) are an inspiration, and his polymathic skill and interdisciplinary breadth continue to make him a model of what a scientist truly should be.

Dr Michael Morrison is a deep wellspring of knowledge and calm, and his expertise in regulatory frameworks, and the social sciences more broadly, were indispensable to this project.

Professor Sam Vanderslott, whose expertise in vaccines and society extend from typhoid in the basement to COVID-19 and beyond, was an outstanding resource, and her generosity in sharing the full resources of the OVG team will always be appreciated.

Professor Sir Andrew Pollard, my shadow supervisor, was always available with blazingly fast and consistently spot-on feedback and encouragement, and he continues to serve as an inspiration of how to be a scientist and a leader.

And of course, Bella and the boys, whose contributions to this project, and my life more generally, are indeed too numerous to count.

Statement of Contribution

This thesis represents entirely my own work. As an integrated thesis, it is based on multiple manuscripts; while my broad interdisciplinary supervisory team provided supervision and feedback, the work is my own.

Dissemination

This is an integrated thesis composed of work that has been published across a number of journals. This includes the following six manuscripts:

Kantor J, Vanderslott S, Morrison M, Pollard AJ, Carlisle RC. The Oxford Needle Experience (ONE) scale: a UK-based and US-based online mixed-methods psychometric development and validation study of an instrument to assess needle fear, attitudes and expectations in the general public. *BMJ Open*. 2023 Dec 13;13(12):e074466. doi: 10.1136/bmjopen-2023-074466. PMID: 38097240; PMCID: PMC10729041.

> This first manuscript outlines the development and validation of a novel scale to assess attitudes to needles, extending the understanding of needle attitudes beyond the existing literature, thus allowing us to more fully define attitudes to the current standard of care (needle and syringe) to which alternatives can then be compared.

Kantor J, Carlisle RC, Morrison M, Pollard AJ, Vanderslott S. Oxford Vaccine Hesitancy Scale (OVHS): a UK-based and US-based online mixed-methods psychometric development and validation study of an instrument to assess vaccine hesitancy. *BMJ Open*. 2024 Oct 9;14(10):e084669. doi: 10.1136/bmjopen-2024-084669. PMID: 39384231; PMCID: PMC11474876.

This second manuscript describes the development and validation of a novel multidimensional vaccine hesitancy scale that can be used, in concert with the prior manuscript, to better understand the nature of vaccine hesitancy and lay the groundwork for de-risking alternative vaccine delivery technologies.

Kantor J, Vanderslott S, Morrison M, Carlisle RC. Discriminative Cut-Offs, Concurrent Criterion Validity, and Test-Retest Reliability of the Oxford Vaccine Hesitancy Scale. *Vaccines*. 2025 Nov 28;13(12):1200. doi: 10.3390/vaccines13121200. PMID: 41441667; PMCID: PMC12737704.

This third manuscript further explores the OVHS, demonstrating that it has good discrimination (it can distinguish between those who refused COVID-19 vaccination and those who did not), that a cut-off value of 35 may be used clinically, and that it has good test-retest reliability when administered to the same participants six weeks apart. This adds significantly to the validity evidence for this scale and enhances its real-world utility.

Kantor J, Carlisle RC, Vanderslott S, Pollard AJ, Morrison M. Development and validation of the Oxford Benchmark Scale for Rating Vaccine Technologies (OBSRVT), a scale for assessing public attitudes to next-generation vaccine delivery technologies. *Hum Vaccin Immunother*. 2025 Dec;21(1):2469994. doi:

10.1080/21645515.2025.2469994. Epub 2025 Mar 3. PMID: 40028861; PMCID: PMC11881882.

This fourth manuscript describes the development and validation of a scale specifically aimed at benchmarking novel vaccine delivery technologies against the standard of care—needle and syringe. Combining these three scales allows researchers and clinicians greater insight into the ways in which emerging technologies are perceived by end-users, patients, and the public.

Kantor J, Morrison M, Vanderslott S, Pollard AJ, Carlisle RC. Public attitudes towards intranasal and ultrasound mediated vaccine delivery: A cross-sectional study in the United Kingdom and United States. *Vaccine*. 2026 Jan 1;69:127950. doi: 10.1016/j.vaccine.2025.127950. Epub 2025 Nov 11. PMID: 41223692.

This fifth manuscript demonstrates the real-world utility of the three novel scales we developed; by combining them with descriptions of next-generation vaccine delivery technologies, we better understand the relative merits of these technologies when both measured against a standard of care (needle and syringe) and against each other. This manuscript demonstrated that ultrasound-mediated vaccination is preferred over both needle and syringe and intranasal vaccination in the United Kingdom and United States.

Kantor J, Morrison M, Vanderslott S, Carlisle RC. En-risk to de-risk: An iterative human factors framework for scale development in emerging technologies. *JAAD Int*. 2025 Dec 4;24:342-345. doi: 10.1016/j.jdin.2025.12.001. PMID: 41567517; PMCID: PMC12818148.

This sixth manuscript underscores the overall value of our quantitative approach, highlighting that converting unmeasurable uncertainty to quantifiable risk is the first step in de-risking biotechnology innovation, that early scale development and deployment can be used to de-risk technologies while providing positive externalities, and proposing the concept of the minimum desirable product—emphasizing public buy-in—as an alternative to the minimum viable product.

Table of Contents

Chapter 1: Literature Review and General Approach 19

1.1 Context.....19

1.2 Instruments for Assessing Needle Fear and Attitudes 22

1.3 Instruments for Assessing Vaccine Hesitancy.....28

1.4 General Approach.....40

1.4.1 Aims.....40

1.4.2 Integrated Thesis and Contributions.....40

1.4.3 The Oxford Needle Experience Scale (ONE scale): Chapter 2 41

1.4.4 Development and Validation of the Oxford Vaccine Hesitancy Scale (OVHS): Chapter 3.42

1.4.5 Discriminative cut-offs, concurrent criterion validity, and test-retest reliability of the Oxford Vaccine Hesitancy Scale (OVHS): Chapter 4.....42

1.4.6 The Oxford Benchmark Scale for Rating Vaccine Technologies (OBSRVT): Chapter 5 42

1.4.7 Public attitudes towards ultrasound-mediated and intranasal vaccination: Chapter 6....43

1.4.8 Towards a Biotechnology De-Risking Toolkit: Chapter 7.....43

1.5 Summary.....43

Chapter 2: The Oxford Needle Experience (ONE) Scale 45

2.1 Introduction 45

2.2 Methods.....46

2.2.1 Domain and Item Generation.....46

2.2.2 Data Collection 50

2.2.3 Domain and Item Reduction and Exploratory Factor Analysis 51

2.2.4 Confirmatory Factor Analysis 54

2.2.5 Reliability and Validity.....56

2.3 Results.....57

2.3.1 Respondent Characteristics.....57

2.3.2 Domain and Item Reduction and Exploratory Factor Analysis 58

2.3.3 Confirmatory Factor Analysis 63

2.3.4 Reliability and Validity.....63

2.4 Discussion 65
2.4.1 Prior instruments 66
2.4.2 Study strengths 67
2.4.3 Study limitations 69
2.4.4 Conclusions 71

Chapter 3: Development and initial validation of the Oxford Vaccine Hesitancy Scale (OVHS) 73

3.1 Introduction 73

3.2 Methods 77
3.2.1 Data Collection and Population 77
3.2.2 Domain and Item Generation 79
3.2.3 Domain and Item Reduction and Exploratory Factor Analysis 80
3.2.4 Confirmatory Factor Analysis 81
3.2.5 Scale Reliability and Validity 81

3.3 Results 82
3.3.1 Respondent Characteristics of Representative Samples 82
3.3.1.1 UK EFA Sample 82
3.3.1.2 US CFA Sample 82
3.3.2 Domain and Item Reduction and Exploratory Factor Analysis 83
3.3.3 Confirmatory Factor Analysis 86
3.3.4 Scale Reliability and Validity 87

3.4 Discussion 88
3.4.1 Vaccine hesitancy 89
3.4.2 Advantages of the OVHS 89
3.4.3 Study strengths 91
3.4.4 Study limitations 92
3.4.5 Conclusions 93

Chapter 4: Discriminative cut-offs, concurrent criterion validity, and test-retest reliability of the Oxford Vaccine Hesitancy Scale 95

4.1 Introduction 95

4.2 Methods 97
4.2.1 Samples 97

4.2.2 Measures 98
4.2.3 ROC curves 99
4.2.4 Cut-off derivation algorithms 99
4.2.5 Test-retest reliability 101

4.3 Results 101
4.3.1 Sample Population 101
4.3.2 ROC Curves and AUC values 102
4.3.3 Cut-Off Derivation 103
4.3.4 Test-retest reliability 107

4.4 Discussion 107
4.4.1 Study strengths 110
4.4.2 Study limitations 110
4.4.3 Conclusions 112

Chapter 5: Development and Validation of the Oxford Benchmark Scale for Rating Vaccine Technologies (OBSRVT) 114

5.1 Introduction 114

5.2 Methods 116
5.2.1 Data Collection, Population, and Overall Approach 116
5.2.2 Domain and Item Generation 118
5.2.3 Domain and Item Reduction and Exploratory Factor Analysis 119
5.2.4 Confirmatory Factor Analysis 119
5.2.5 Reliability and Validity 120

5.3 Results 120
5.3.1 Characteristics of the Representative Samples 120
5.3.1.1 UK EFA Sample 120
5.3.1.2 US CFA Sample 121
5.3.2 Domain and Item Reduction and Exploratory Factor Analysis 121
5.3.3 Domains Included in the Final Scale 122
5.3.4 Confirmatory Factor Analysis 125
5.3.5 Scale Reliability and Validity 125

5.4 Discussion 126
5.4.1 Study strengths 127
5.4.2 Study limitations 129
5.4.3 Conclusions 131

Chapter 6: Public attitudes towards intranasal and ultrasound-mediated vaccine delivery: A cross-sectional study in the United Kingdom and United States 132

6.1 Introduction 132

6.2 Methods 135
6.2.1 Population, Sample Size, and Framing 135
6.2.2 Validated Scales 137
6.2.3 Willingness to Pay and Stated Preferences 138
6.2.4 Discrete Choice Experiment 139
6.2.5 Associations with Needle and Vaccine Hesitancy 139
6.2.6 Statistical Analysis 140

6.3 Results 141
6.3.1 Samples 141
6.3.2 UK Sample 141
6.3.3 US Sample 141
6.3.4 Outcome Measures 142
6.3.4.1 OBSRVT 142
6.3.4.2 Willingness to Pay 143
6.3.4.3 Stated Preferences 145
6.3.4.4 Discrete Choice Experiment 145
6.3.5 Associations with Needle and Vaccine Hesitancy 146

6.4 Discussion 153
6.4.1 Study strengths 156
6.4.2 Study limitations 157
6.4.3 Conclusions 159

Chapter 7: Conclusions: Developing the minimum desirable product by de-risking biotechnology ventures through early stage validated scale development and iterative human factors research 161

7.1 Human factors research 161

7.2 From uncertainty to risk: En-risk to de-risk 163

7.3 Instrument and Scale based risk quantification 163
7.4 Summary: The ultrasound-mediated vaccine delivery model 164
7.5 From the MVP to the MDP 165
7.6 Future directions 167
7.7 Conclusions 167
References 169
Appendix: Glossary of Psychometric Terms 200

List of Figures

Figure 1.1 Psychometric approaches used to develop existing vaccine hesitancy scales. 39

Figure 1.2 Standardized steps in scale development..... 41

Figure 4.1 ROC curve for the OVHS in the $UK_{1.1}$ dataset..... 103

Figure 4.2 ROC curve for the OVHS in the $UK_{1.2}$ dataset..... 103

Figure 4.3 Heatmap demonstrating the near-optimal plateau..... 104

Figure 4.4 Two-graph ROC-style figure. 106

Figure 6.1 Demand curve illustrating the relationship between price and demand. . 144

Figure 7.1 Technology readiness level (TRL) and public acceptance of vaccine delivery technologies. All values are demonstrative. 162

Figure 7.2 Traditional research and development (R&D) focuses on the minimum viable product (MVP), while incorporating human factors research early allows for a focus on the minimum desirable product (MDP). 165

List of Tables

Table 1.1 Preexisting needle fear scales. 23

Table 1.2 Summary of preexisting vaccine hesitancy scales. 30

Table 2.1 Factor loadings for the ONE exploratory factor analysis (EFA). 60

Table 2.2 Factor intercorrelations in the UK EFA. 61

Table 2.3 The ONE Scale. All answer choices are rated using a 5-point Likert scale (strongly disagree to strongly agree). For questions 1-16, strongly agree is scored as 5, while for questions 17-19 strongly agree is scored as one (reverse scoring). Higher values reflect more hesitancy regarding needle use. The total possible ONE score ranges from 19 to 95. 62

Table 2.4 Median and interquartile ranges for ONE sub-scales and the overall scale. 63

Table 2.5 Subscale and overall reliability, as measured by McDonald's Omega. 64

Table 3.1 Rotated factor loadings for the OVHS. 85

Table 3.2 Factor intercorrelations. 85

Table 3.3 The OVHS. 86

Table 3.4 Comparison of scores between UK and US Samples. 87

Table 3.5 Subscale and overall reliability (assessed using McDonald's Omega). 87

Table 4.1 Discriminative performance (ROC AUC) of the OVHS for classifying historical COVID-19 vaccination refusal across samples. 102

Table 4.2 Index of Union (IU) across candidate OVHS cut-offs (34–38) by dataset, with per-dataset IU minimum and aggregate IU included. 104

Table 4.3 Discriminative performance of total OVHS score≥35 for classifying self-reported historical COVID-19 vaccine refusal across five samples. 105
Table 4.4 Operating characteristics of sample-specific optimal OVHS cut-offs based on maximizing Youden's J for each sample. 106
Table 4.5 Two-way mixed-effects absolute-agreement single- and average-measure ICC(A,1) and ICC(A,2) model results, with n=462 over two timepoints six weeks apart in a UK sample. 107
Table 5.1 Rotated factor loadings for the OBSRVT. 123
Table 5.2 The OBSRVT scale. 124
Table 5.3 Subscale and overall reliability (assessed using Cronbach's alpha). 125
Table 5.4 Representative comparison of scores for the OBSRVT scale when assessing a nasal spray vaccine between UK and US Samples. 128
Table 6.1 Overall and subscale OBSRVT scores for nasal spray and ultrasound-mediated vaccination. 143
Table 6.2 Willingness to pay for various delivery technologies, by sample. 144
Table 6.3 Stated preferences (0-100, with higher values representing more desirability) for vaccine delivery technologies by sample. 145
Table 6.4 OBSRVT scores stratified by ONE scale quartile. 147
Table 6.5 OBSRVT score by OVHS scale quartile. 148
Table 6.6 Mean willingness to pay (£/$) by ONE scale quartile. 149
Table 6.7 Willingness to pay (in £ or $, as appropriate) by OVHS quartile. 151
Table 6.8 Stated preferences (0-100, with higher values representing more desirability) by ONE scale quartile. 151

Table 6.9 Stated preferences (0-100, with higher values representing more enthusiasm) by OVHS quartile. 153

Chapter 1: Literature Review and General Approach

1.1 Context

Vaccine hesitancy has been identified by the World Health Organization (WHO) as one of the 10 greatest threats to global health, with its urgency as a global health priority underscored by high rates of COVID-19 vaccine refusal, with over 20% of eligible adults ultimately going unvaccinated (Lazarus et al, 2020; Mathieu et al, 2021; World Health Organization, 2019). Given the importance of behavioural and attitudinal considerations in modulating vaccine acceptance (Kantor, 2020), even subtle changes in methods of vaccine delivery, such as avoiding the use of a needle and syringe, may decrease vaccine hesitancy, leading to significant global health benefits.

Multiple factors are associated with vaccine hesitancy, including mistrust of healthcare providers, concerns regarding government control, fear of needles, and worries about the long- and short-term health risks of vaccination (Marti et al, 2017). While anti-vaccine groups tout safety concerns as their motivation, this represents only one of many putative contributors worldwide (Dubé et al, 2014), and confidence in the importance of vaccination may ultimately be more important than safety concerns (de Figueiredo et al, 2020). Vaccine hesitancy also cannot be legislated away, as hesitancy spurred by single-vaccine mandates may metastasize to vaccination programmes more generally, while mandates and vaccine passports may decrease the likelihood of certain groups accepting vaccination (Bardosh et al, 2022; de Figueiredo et al, 2021; Durbach, 2000; Eagan et al, 2023; Johnson et al, 2020; Ward et al, 2022).

Given the multifactorial drivers of vaccine hesitancy, as well as regional and temporal variation (de Figueiredo & Larson, 2021), developing validated and reliable instruments to assess this challenge is an international public health priority (Salmon et al, 2015a).

Numerous alternatives to traditional needle and syringe-based vaccination have been proposed, including jet and ballistic injectors, microneedles, DNA tattooing, iontophoresis, and ultrasound (cavitation-mediated) vaccine delivery (Hettinga & Carlisle, 2020). Since at least 16% of adults refuse yearly influenza vaccination due to a fear of needles and 27% of hospital employees express a fear of needles, alternatives to needle and syringe-based vaccination are needed (McLenon & Rogers, 2019), particularly since needle-phobic parents are more likely to delay their children's vaccinations (Callaghan et al, 2019).

Vaccine delivery to the highly immunologically active dermis, rather than muscle, may lead to not only increased vaccine efficacy, but also to a decreased total dose requirement (Anand et al, 2017; Arnou et al, 2010; Okayasu et al, 2017). Since vaccine hesitancy may be partially driven by concerns regarding dosage (Callaghan et al, 2019; Marti et al, 2017), this represents another potential opportunity to improve vaccine uptake.

While a wide range of alternatives to needle and syringe-based vaccination have been proposed, one advantage of ultrasound-mediated vaccine delivery is the possibility that the cavitation event itself could act as an adjuvant (Seah & Teo, 2018). If so, ultrasound-mediated vaccine delivery could provide four important advantages over needle and syringe-based vaccination from a public acceptability standpoint:

superficial vaccine placement, avoidance of needles, decreased dose requirement, and decreased adjuvant requirement. The relative importance of these putative advantages in encouraging public acceptance, however, has not been explored. Since framing affects acceptability and valuation of medical treatment choice (McNeil et al, 1982; Tversky & Kahneman, 1981), there may be significant opportunities for increasing vaccine uptake and reducing hesitancy by developing an improved understanding of the individual drivers of vaccine hesitancy and engineering alternative vaccine delivery systems accordingly. This is particularly important as ultrasound is viewed favourably in the healthcare context, suggesting that it may hold significant promise in encouraging vaccination (Andersen et al, 2021; Hettinga & Carlisle, 2020).

To rigorously assess the relative merits of alternatives to needle and syringe-based vaccination, it is helpful to understand the motivations behind needle fear and needle hesitancy, and – more importantly – to reliably *quantify* the degree of needle hesitancy in a population of interest. Similarly, understanding the relative contribution of various factors (such as needle fear, scientific mistrust, or safety concerns) in vaccine hesitancy more generally – and, again, quantifying this hesitancy – is a central step in developing alternative vaccine delivery approaches. Indeed, it is essential that the development of new alternative technologies is guided by an understanding of the aspects of current practice which drive hesitancy, as otherwise such new technologies may be similarly avoided. This literature review therefore focuses on the literature describing a) instruments and scales for assessing the fear of needles and b) instruments and scales for measuring vaccine hesitancy.

1.2 Instruments for Assessing Needle Fear and Attitudes

Several instruments have been developed to assess fear of needles in the past. A summary of the scales discussed below is included in Table 1.1. Psychometric term definitions are included in the Appendix.

Scale	Full Name	Citation	Items	Domains	Notes
IPS	Injection Phobia Scale-Anxiety	Öst et al, 1992; Olatunji et al, 2010	18	2 (contact fear, distal fear)	Developed on 40 phobia outpatients; not validated until 2007; not designed for general population
MFS	Medical Fear Survey	Kleinknecht et al, 1996, 1999	50	5 (injections, sharp objects, blood, mutilation, examinations + symptoms	Very lengthy; includes non-needle domains (mutilation, examinations); US university students only
MFS-Short	Medical Fear Survey–Short Version	Olatunji et al, 2012	25	5 (injections, sharp objects, blood, mutilation, examinations + symptoms)	Includes broad range of non-needle domains; US/Dutch university students
BISS	Blood-Injection Symptom Scale	Page et al, 1997	17	3 (faintness, anxiety, tension)	Binary responses; focus on physical symptoms only; does not address psychological/historical contributors
BIF	Blood/Injection Fear Scale	Kose & Mandiracioglu, 2007	20	2 (injection, blood)	EFA only; no goodness of fit statistics; Turkish hospital population limits generalizability
MBPI	Multidimensional Blood/Injury Phobia Inventory	Wenzel & Holt, 2003; van Overveld et al, 2011	40	6 (injections, hospitals, fainting, blood-self, injury, blood-others)	Designed for pathological reactions; no qualitative methodology; literature-derived domains only
SPQ-BII	Specific Phobia Questionnaire (BII subscale)	Ovanessian et al, 2019	10	Not explored	Part of broader phobia scale; developed on anxiety disorder patients; correlates

					with other phobia subtypes
MFIS	Multidimensional Fear of Injection Scale	Hako et al, 2022	20	4 (direct fear, indirect fear, physiological response, avoidance)	No qualitative methodology; theory-driven; excludes past experiences domain; Japanese university students only

Table 1.1 Preexisting needle fear scales.

The notes column reflects issues relevant when developing a scale aimed at understanding needle fear generally and does not represent actual limitations of preexisting scales.

The 18-item Injection Phobia Scale (IPS) was developed on a sample of 59 subjects seen as outpatients in a Swedish mental hospital and included separate subscales for anxiety and avoidance (Öst et al, 1992). The IPS did not report construct or criterion validity—indeed, the first validation study was performed in 2007 (Olatunji et al, 2007b)—and the only test statistic presented by the authors was internal consistency, as measured by Cronbach's alpha, which was 0.86 for the anxiety subscale and 0.80 for the avoidance subscale. Little information was provided regarding the methodology used to create the items, and the factorial and domain structure of the IPS was not explored. Still, this scale has several important advantages, including the inclusion of a wide range of items, development specifically to evaluate injection phobia, ability to capture both the anxiety and avoidance aspects of needle phobia, and sensitivity to change as demonstrated by improvement in scores in response to therapeutic intervention. Yet though the IPS represents an important step, its generalizability is limited by its narrow sample population. Therefore, since the small number of subjects included in IPS development were all previously diagnosed with true needle phobia, this scale was not developed as an instrument for use on the general population.

The IPS has been broadly adopted, has been used to assess needle phobia in over 125 manuscripts, and has been validated in other languages (Bani et al, 2020). Moreover, as highlighted by others (Olatunji et al, 2010), it has shown discriminant validity in its potential to distinguish between injection phobia and other phobias (Sawchuk et al, 2002). Finally, it has also demonstrated adequate convergent validity in its association with disgust (Olatunji et al, 2007a) and contamination (Sawchuk et al, 2000).

Decades after its initial development, the factorial structure of the IPS was assessed by performing an exploratory factor analysis on a population of 498 university undergraduates in the United States, followed by a confirmatory factor analysis on an additional 567 students from the same university population (Olatunji et al, 2010). This approach allows the observed variance in correlated test items to be attributed to a smaller number of unobserved latent factors, and the authors demonstrated a two-factor structure, encompassing contact fear (fear relating to actual needle injection) and distal fears (fear of watching or discussing injections). Several other studies were performed that further bolster the value of the IPS (van Overveld et al, 2006; Watson et al, 1988).

Several other scales have been developed to assess injection fear. The 50-item Medical Fear Survey (MFS) examines an array of fears relating to medical procedures, including injections as well as blood, physical examination, and mutilation (Kleinknecht et al, 1999; Kleinknecht et al, 1996). The MFS was developed based on the Mutilation Questionnaire, a scale that included true/false responses to a range of questions assessing blood, injury, injection, and mutilation fears (Klorman et al, 1974).

The factorial structure of the MFS was evaluated in a group of 934 university undergraduates in the United States, and three of the seven factors—fear of the sight of blood, needles and injections, and physical symptoms—are directly relevant to understanding needle fear. These factors were also associated with a history of fainting in the past ($p<0.001$). While the MFS may therefore be a useful adjunct in evaluating injection fears, its length and inclusion of multiple domains that do not relate to needle fear, such as domains relating to fear of mutilation, deformity, and examinations, may make it less useful when trying to understand needle fear as a construct.

An abbreviated 25-item version of the MFS was later developed on 931 university students in the United States (Olatunji et al, 2012). The authors relied on a combination of Mokken scaling analysis, a non-parametric technique to evaluate clustering between items, and visual inspection of scale items to determine whether items were candidates for deletion (Mokken & Lewis, 1982). Following an item-response theory approach, they also evaluated items for essential independence (Stout, 1990), yielding a refined 25-item scale. Of note, the authors performed a cross-cultural study by including an additional set of 258 Dutch students in their analyses, and found that responses did not differ substantially between American and Dutch students. Additional studies included classification validity and structural validity, highlighting the adequate performance of the MFS-short version in evaluating a range of medical fears, including needle fear (Stout, 1990).

The Blood-Injection Symptom Scale (BISS), which assesses domains of faintness, anxiety, and tension, and was developed on a population of 605 Australian

university students followed by a group of 80 community members self-identified as having pre-existing fears relating to injections (Page et al, 1997a). This scale was expressly developed to distinguish between those who primarily exhibited fear, and therefore concerns regarding injections, and those who exhibited fainting, and therefore concerns regarding blood (Ost, 1992). The 17-item scale was scored with binary responses and included items relating exclusively to past experiences when undergoing injections, such as whether respondents had previously experienced dizziness, sweating, trembling, or blurred vision when undergoing injections. Therefore, while this scale effectively achieves its stated purpose—quantifying the *physical* symptoms associated with physiological and psychological responses to injections—its narrow focus does not allow it to address the various psychological and historical contributors to needle fear that are critical for a multidimensional scale that explores responses to needles.

The blood/injection fear scale was developed on a Turkish population of 1,500 patients attending a hospital for examination or blood donation, and is a 20-question scale that includes two domains, injection and blood (Kose & Mandiracioglu, 2007). The injection domain includes items that largely focus on fear, while the blood domain items largely address disgust at the sight of blood. When tested for internal consistency, or how closely related a set of items are as a group, a Cronbach's alpha value of 0.98 was obtained, suggesting a high level of reliability (Tavakol & Dennick, 2011). As the study included only an exploratory factor analysis, no goodness of fit statistics were available for assessment, and its generalizability may be limited due to the population used for scale development.

The 40-item Multidimensional Blood/ Injury Phobia Inventory (MBPI) was developed on a US university student population of 558 individuals, with a follow-up study on 616 students and university employees, to assess phobias, and was designed to assess pathological reactions to blood and needles (van Overveld et al, 2011; Wenzel & Holt, 2003). Domains, including injections, hospitals, fainting, blood (self), injury, and blood (others) were determined based on the literature, and a qualitative analysis was not performed to broaden the range of domains.

Blood-injection-injury was also assessed as part of a 10-item subscale of the Specific Phobia Questionnaire, a lengthy scale addressing phobias ranging from fear of animals to the natural environment; this scale was developed on a cohort of 1,031 patients with known anxiety disorders and a smaller group of 182 Canadian university students (Ovanessian et al, 2019). The convergent validity of this scale was adequate ($r=0.44$, $p<0.001$) when measured against the BISS, though it also correlated with situational fears and fears of the natural environment ($r=0.58$, $p<0.001$). This scale may therefore be of limited use when attempting to understand needle fear and expectations in the broader population.

Finally, the multidimensional fear of injection scale was developed on 419 Japanese university students to evaluate fear of injections exclusively, encompassing domains of direct fear, indirect fear, physiological response, and avoidance (Hako et al, 2022). While multidimensional in that it was designed to address multiple contributors to fear, this scale relied on a theoretical model that phobia is driven by direct and indirect fear, physiological response, and avoidance. This work focused exclusively on fear and relied on a population of Japanese university students, where

generalizability may be limited compared with using quota sampling. While the authors acknowledge the importance of negative historical experiences in developing fear of needles, their scale does not include this as a separate domain, and no items in the scale address past negative experiences.

Almost all of these scales focus on addressing needle *phobia*; while it is helpful to identify patients with a true phobia, this represents a very small subset of the population—in the range of 2-3% (Bienvenu & Eaton, 1998; Deacon & Abramowitz, 2006; Fredrikson et al, 1996)—and therefore expanding the definition to capture needle fear more broadly may allow public health authorities to address the concerns of a larger portion of the population. Moreover, those with a true phobia of needles are readily able to self-identify, and so developing of a scale to identify this phobia may add little to what could likely be elicited with a single-question survey. In contrast, understanding the multidimensional nature of needle attitudes and fear allows for a more nuanced appreciation of the range of responses to needle injection, and may ultimately provide a more clinically meaningful and useful approach in designing alternatives to needle and syringe-based vaccination.

1.3 Instruments for Assessing Vaccine Hesitancy

Vaccine hesitancy is a broad construct, and was defined by a WHO Strategic Advisory Group of Experts (SAGE) working group as 'a delay in acceptance or refusal of vaccination despite availability of vaccination services [and a phenomenon that is] complex and context specific, varying across time, place, and vaccines [and] influenced by factors such as complacency, convenience and confidence' (MacDonald

& Sage Working Group on Vaccine Hesitancy, 2015). This definition was later revised as 'a motivational state of being conflicted about, or opposed to, getting vaccinated; this includes intentions and willingness' (World Health Organization, 2022).

A systematic review of 422 studies addressing vaccine hesitancy highlighted the complexity of the concept, noting that three central approaches to understanding the phenomenon are reflected in the literature: cognition or affect, behaviour, and decision making. The authors suggested that the central theme of vaccine hesitancy is a 'state of indecisiveness' regarding the decision to vaccinate (Bussink-Voorend et al, 2022). Several validated scales have been developed to evaluate vaccine hesitancy, and a scoping review has examined the past decade of scale development in this field (Oduwole et al, 2022). A summary of the scales discussed below is included in Table 1.2.

Scale	Full Name	Citation	Items	Domains	Notes
PACV	Parent Attitudes about Childhood Vaccines	Opel et al, 2011a, 2011b	15 (from 18)	3 (general attitudes, safety & efficacy, behaviour)	Paediatric vaccines only; limited multidimensionality
SAGE Survey	SAGE Vaccine Hesitancy Working Group Survey	Larson et al, 2015	10	Not formally explored (content addresses effectiveness, risk, and trust)	No factor analysis; no psychometric testing
VCBS	Vaccine Conspiracy Belief Scale	Shapiro et al, 2016	7	1 (unidimensional; items address trust in vaccine manufacturers and government)	Limited to conspiracy beliefs only
VCS	Vaccine Confidence Scale	Gilkey et al, 2014, 2016	8 (4-item short form)	3 (benefits, harms, trust)	Modest number of items and domains
VAI	Vaccine Acceptance Instrument	Sarathchandra et al, 2018	20 (10-item short form)	5 a priori (safety, effectiveness, acceptance of scheduling, positive beliefs, legitimacy of mandates); PCA suggested 3-factor	Small sample; rejected PCA findings

VCI	Vaccine Confidence Index	Frew et al, 2019	8	5 (information environment, trust, healthcare provider, attitudes and beliefs, social norms)	Five-factor unlikely with eight items; unstable
VAX	Vaccine Attitudes Examination	Martin & Petrie, 2017b	12	4 (mistrust of vaccine benefits, worries about unforeseen effects, concerns about commercial profiteering, preference for natural immunity)	Non-representative; small focus groups
VHS	Vaccine Hesitancy Scale	Shapiro et al, 2018	9	2 (lack of confidence, risks)	Simple two-domain; only two risk items; paediatric
5C	5C Scale (Psychological Antecedents)	Betsch et al, 2018	15 (5-item short form)	5 (confidence, complacency, convenience, calculation, collective responsibility)	A priori framework may limit real-world capture
MVHS	Multidimensional Vaccine Hesitancy Scale	Howard, 2022	32	8 (health risks, cost, physical pain, inconvenience, personal reactions, access, self-perception of health, forgetting)	Non-representative; no trust items; redundant items

Table 1.2 Summary of preexisting vaccine hesitancy scales.

The notes column reflects issues relevant when developing a multidimensional scale aimed at understanding vaccine hesitancy generally and does not represent actual limitations of preexisting scales.

The Parent Attitudes about Childhood Vaccines (PACV) scale, the most highly cited scale in evaluating vaccine hesitancy (Bussink-Voorend et al, 2022), was developed to identify parents who are vaccine hesitant (Opel et al, 2011a; Opel et al, 2011b). The initial validation study explored the face and content validity of the PACV using the original 18-item scale with domains encompassing immunization behaviour, beliefs about safety and efficacy, attitudes about mandates and exemptions, and trust. A separate study was then performed that reduced the number of items in the scale to 15 and suggested a three-factor solution with domains including general

attitudes, safety & efficacy, and behaviour (Opel et al, 2011b). Importantly, the PACV addresses parental attitudes rather than attitudes of an individual regarding their own vaccination, and given the centrality of the paediatric population to vaccination efforts worldwide, new technologies may well be first deployed in children. Still, while the PACV has important strengths, including its parsimony and multidimensional approach, it was developed exclusively for childhood vaccinations and therefore does not capture the rich multidimensionality of contemporary vaccine hesitancy across other groups.

Survey instruments were also developed by the SAGE Vaccine Hesitancy Working Group, ultimately yielding three separate sets of items: core questions, Likert-scale questions, and open-ended questions (Larson et al, 2015). The core questions include binary items that directly ask about beliefs regarding the effectiveness of vaccines and a history of reluctance to vaccinate or vaccine refusal, while the Likert-scale questions represent a 10-item set that addresses beliefs regarding the effectiveness of vaccines, risk, and trust. While these questions address a range of components that align well with the original SAGE definition of vaccine hesitancy, the factor structure of these questions was not explored and psychometric testing was not performed.

Conspiracy beliefs increasingly represent an important element of vaccine hesitancy, and the Vaccine Conspiracy Belief Scale (VCBS) was developed specifically to address these concerns (Shapiro et al, 2016). The VCBS, developed on a group of 1,427 Canadian parents, demonstrated good convergent validity, as scores were inversely related to willingness to vaccinate a child against human papillomavirus

(Bruder et al, 2013). Factor analysis demonstrated a unidimensional solution. The seven items relate largely to trust in vaccine manufacturers and the government, though their comprehensiveness was limited by perceptions of common conspiracy theories regarding vaccines that were prominent in 2016, and the scale does not address vaccine hesitancy more broadly.

The Vaccine Confidence Scale (VCS) was developed using data from almost 10,000 parents who completed a nationally representative telephone survey in the United States (Gilkey et al, 2014; Gilkey et al, 2016). The VCS includes three domains (benefits, harms, and trust) and eight items, and the authors found that VCS scores were associated with vaccine refusal. A 4-item short form of the questionnaire was also evaluated. While the VCS was developed on a large nationally representative population, the relatively modest number of items and domains means that it may not capture many of the complex and dynamic contributors to contemporary vaccine hesitancy.

The majority of scales used to assess vaccine hesitancy were designed for administration to parents; the Vaccine Acceptance Instrument (VAI) was developed as a multidimensional scale to overcome this limitation (Sarathchandra et al, 2018). The VAI is a 20-item scale that examines five domains: safety, effectiveness, acceptance of scheduling, positive beliefs in vaccines, and the perceived legitimacy of those mandating vaccinations. A 10-item short form was also validated by the authors. This scale has several strengths, including the inclusion of reverse scored items (to limit acquiescence bias) and the use of a balanced questioned stem, as well as equal numbers of items across domains. The theoretically sound domains addressed by the

authors were not reflected by their principal components analysis, which instead suggested a three-factor solution. While this discrepancy may have resulted from an inadequate sample size (246 participants in total), the authors chose to overlook the results of their statistical analysis in favour of maintaining their preferred 5-factor solution based on 'the conceptual and theoretical justifications for distinguishing the five dimensions', and therefore the statistical underpinnings of their approach—and the broad applicability of the scale—may be limited.

The Vaccine Confidence Index (VCI) was developed specifically to assess parental trust in the US vaccination system (Frew et al, 2019). The scale was developed on 1,160 participants and the five a priori domains (information environment, trust, healthcare provider, attitudes and beliefs, and social norms) were based on theoretical categories highlighted by the US National Vaccine Advisory Committee (Orenstein et al, 2015). An exploratory factor analysis was used to refine the initial item set, though little methodological detail was provided regarding this process. The resulting eight-item scale appears to capture a single underlying construct, vaccine confidence, rather than preserving the original five-domain structure. The authors also validated the VCI against self-reported childhood vaccination status, and compare the VCI with existing instruments including the PACV, highlighting differences in target populations, dimensionality, and domains assessed.

Among the only scales to address concerns regarding perceived pharmaceutical company profiteering, the Vaccine Attitudes Examination (VAX) scale is a 12-item four-domain instrument that addresses mistrust of vaccine benefits, worries regarding future unforeseen effects, concerns about commercial profiteering,

and a preference for natural immunity (Martin & Petrie, 2017b). Development was first performed through focus groups in doctors' surgeries and ultimately included 501 student and convenience-sample participants in the US. The authors found that the VAX scale score correlated with past vaccination behaviour, its test-retest reliability (measured one-month apart) was adequate, and convergent validity (with the PACV) was adequate. Strengths of this study include its broad approach to domain generation and the use of statistical approaches such as polychoric correlation matrices, which are most appropriate for Likert-scale data (Flora & Curran, 2004; Holgado–Tello et al, 2010). Limitations include the limited domain range, the use of university students and a non-representative convenience sample (only 2% of community respondents had a university degree) and the small number (n=14 total) of participants included in the domain- and item-generation focus groups. The authors' approach to assessing prior vaccination behaviour may also be problematic, as two of the three dichotomous items addressed prior-year influenza vaccination—a vaccine that is generally perceived as safe (Lutz et al, 2020) but where uptake hovers around 50% and where even mild prior-year side effects have been associated with vaccine refusal (Smith et al, 2017).

The Vaccine Hesitancy Scale (VHS) was developed by a SAGE working group on a set of 3,779 Canadian parents and included nine items over two domains: lack of confidence and risks (Shapiro et al, 2018). Adequate convergent validity was seen with the VCBS (Shapiro et al, 2016), and criterion validity was demonstrated as each subscale was associated with a history of vaccine refusal ($p<0.001$). Of the nine items, seven loaded onto the lack of confidence domain and two onto the risk domain, and

Cronbach's alpha for the two domains was 0.92 and 0.64, respectively. This impressive parsimony of the VHS is offset by its simple 2-domain solution, a limitation that is particularly acute given that only two items loaded on the risk domain. Thus the many domains that likely contribute to vaccine hesitancy remain largely unexplored by this scale.

The original SAGE definition of vaccine hesitancy focused on the three 'C' features of this phenomenon: confidence, complacency, and convenience (MacDonald & Sage Working Group on Vaccine Hesitancy, 2015). To expand further on these phenomena, the 5C scale was developed to also include calculation and collective responsibility; the authors theorized that if they could evaluate the 'psychological antecedents to vaccination' (Brewer et al, 2017), this would allow them to better understand the complex phenomenon of vaccine hesitancy (Betsch et al, 2018). The authors relied initially on a convenience sample of 1,445 German residents to assess the scale's validity followed by a 1,003-participant representative German sample to assess whether the scale responded better when used to evaluate responses to specific vaccines, and both a 15-item and a 5-item scale were developed and validated by the authors. The central strength of this scale is its multidimensional approach that evaluates vaccine hesitancy more broadly than prior hesitancy scales, which focus largely on confidence alone, though its reliance on an a priori derived theoretical framework to explain the behavioural antecedents of vaccine hesitancy may limit the ways in which it captures the full range of domains contributing to hesitancy in the real world.

The Multidimensional Vaccine Hesitancy Scale (MVHS), a 32-item eight-domain scale that was developed on a survey-panel population, was created in an attempt to broaden the number of domains included in existing scales (Howard, 2022). Domains included health risks; cost; physical pain; inconvenience; personal reactions; access; self-perception of health; and forgetting. Thus, a central strength of the MVHS is its broad domain structure. Limitations, however, include the broad conceptualization of vaccine hesitancy as 'having concerns about vaccines, regardless of actual vaccine receipt', which arguably would limit the utility of such a scale in a public health setting where actionable data are used to create interventions, the use of non-representative samples across varying populations, and the use of statistics assuming continuous variables when assessing Likert-scale data. Some of the findings are also counterintuitive, as higher access and inconvenience scores, which suggest difficulty obtaining a vaccine due to availability and logistic challenges, are associated with *increased* vaccine adoption. The author argues that this is because such 'negative perceptions' may not deter vaccine use, though this still does not explain the increase in vaccine usage, and perhaps a better explanation is that paradoxically those who are most enthusiastic about vaccines are likely the population most *concerned* about access, and therefore that concerns about access perhaps should not be a focus when trying to capture those who are hesitant about vaccines. The MVHS also approaches the cost of vaccines as a concept similar to access ('vaccines cost too much') rather than addressing concerns regarding profiteering. The author also over-generalizes from a series of small studies to make broad claims regarding understanding vaccine hesitancy, arguing that since health risks and efficacy doubts load on the same factor

they should be seen as occupying the same criterion space; in fact, such loadings are highly dependent on the factorial structure of the model and the number of factors selected for retention, and therefore such conclusions are unwarranted. This over-reliance on statistical findings in the absence of theoretical justification is also reflected in the lack of any trust- or conspiracy-related domains or items in the final scale; given the overarching importance of mistrust in vaccine manufacturers and the government in driving the present wave of vaccine hesitancy, this is an important omission. Perhaps the greatest limitation of the MVHS is that many of the items are redundant and, as acknowledged by the author, 'near synonyms'. For example, under cost, the four distinct items are 'vaccines cost too much', 'I am unable to get vaccines because they cost too much', vaccines are too expensive', and 'without health insurance vaccines cost too much', all essentially synonymous.

While multiple other scales have been developed that evaluate subsets of vaccine hesitancy such as medical mistrust (Sanford & Clifton, 2022) and others were designed to be disease-specific such as those for the COVID-19 (Kotta et al, 2022), measles, mumps, and rubella vaccines (Brown et al, 2011), or Mpox (Gao et al, 2024), such scales may be of limited utility when evaluating vaccine hesitancy more broadly. Similarly, many pre-existing vaccine hesitancy scales were designed for paediatric vaccines exclusively (Opel et al, 2011a; Opel et al, 2011b; Sarathchandra et al, 2018; Shapiro et al, 2018; Yao et al, 2023) and are therefore less appropriate for understanding responses to vaccination more generally, while other scales demonstrate a lack of appreciation of critical concepts such as trust (Larson et al, 2018).

As highlighted by others (Betsch et al, 2018), most pre-existing vaccine hesitancy scales focus exclusively on a narrow range of domains, such as confidence, and even those that depart from this tend to rely on theoretical models—rather than qualitative studies eliciting public opinion—in developing a range of domains for inclusion (Gilkey et al, 2014). As noted above, a range of psychometric approaches were used in developing these scales (Figure 1.1). The VAX scale was developed using qualitative data from small focus groups, but the small number of domains and the non-representative population coupled with the small sample size for confirmatory factor analysis—as well as the limited reporting of model fit indices (only the Normed Fit Index and Tucker Lewis Index were reported)—are important limitations. Similarly, the MVHS was developed based on qualitative data and included a broader set of domains but excludes trust and conspiracy theories entirely. Therefore, a new broad multidimensional scale is needed to better capture the range of experiences and expectations in the contemporary vaccine hesitancy landscape.

Scale	Internal Consistency Reliability	Exploratory Factor Analysis	Confirmatory Factor Analysis	Convergent / Discriminant Validity	Criterion Validity	Test-Retest Reliability
PACV	✓	✓	×	×	✓	✓
SAGE Survey	×	×	×	×	×	×
VCBS	×	✓	×	✓	✓	×
VCS	✓	✓	✓	×	✓	×
VAI	✓	~	×	×	×	×
VCI	✓	✓	×	×	✓	×
VAX	×	✓	~	✓	✓	✓
VHS	✓	✓	×	✓	✓	~
5C	✓	✓	×	✓	~	~
MVHS	✓	✓	✓	✓	×	×

✓	Property reported
×	Property not reported
~	Partial, ambiguous, or limited evidence

Figure 1.1 Psychometric approaches used to develop existing vaccine hesitancy scales.

1.4 General Approach

1.4.1 Aims

Developing and validating survey instruments to assess public acceptance of alternatives to needle and syringe-based vaccination is an essential and helpful first step in encouraging adoption of such alternative technologies. Indeed, such survey instruments may better inform outreach and education efforts globally while strategically developing alternative vaccine delivery approaches that not only improve the immunologic effectiveness of vaccines but increase public willingness to accept them as well. This approach is also consistent with the Responsible Research and Innovation (RRI) scheme for scientific and technological development, which highlights that the very act of public engagement may itself improve acceptance of vaccines (Shelley-Egan et al, 2018).

We therefore developed a validated needle attitude scale; a validated scale assessing vaccine hesitancy; a benchmark scale to assess attitudes to alternatives to needle and syringe-based vaccination; and assessed the relative valuation of the putative advantages of intranasal and ultrasound-mediated vaccination in both the UK and US general population.

1.4.2 Integrated Thesis and Contributions

This is an integrated thesis based on six separate manuscripts published or submitted for publication, with each manuscript representing a separate chapter; a brief outline of each study is included below. Given that this thesis is designed so that it represents a cohesive whole but also that each chapter can be read independently,

some redundancy in abbreviations, methods, and discussion between chapters is purposeful and expected. This thesis represents entirely my own work in collaboration with feedback from my supervisor co-authors; these collaborative contributions included supervisory oversight and feedback on manuscript drafts.

1.4.3 The Oxford Needle Experience Scale (ONE scale): Chapter 2

First, to better understand how needle attitudes interact with choices regarding vaccination techniques, we developed and validated the Oxford Needle Experience (ONE) scale. We adopted an iterative, standardized methodology to develop and validate this scale. In brief, after a comprehensive literature review, a series of online surveys were conducted to aid in item generation and domain identification, followed by an iterative process to establish face validity, a robust maximum likelihood exploratory factor analysis, and a confirmatory factor analysis; this approach is outlined in Figure 1.2.

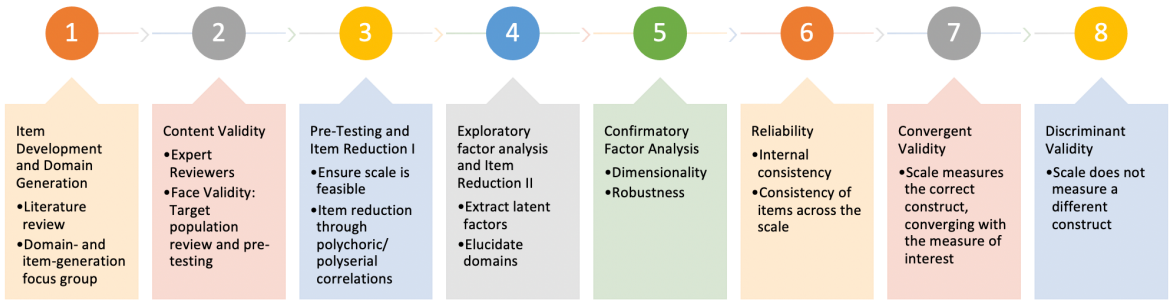


Figure 1.2 Standardized steps in scale development.

1.4.4 Development and Validation of the Oxford Vaccine Hesitancy Scale (OVHS): Chapter 3

Second, we developed and validated the Oxford Vaccine Hesitancy Scale (OVHS), a novel multidimensional scale assessing vaccine hesitancy, following a similar structured approach. While vaccine hesitancy is a heterogeneous phenomenon, better understanding the domains that contribute to this phenomenon may help design both behavioural interventions and alternative vaccine delivery technologies.

1.4.5 Discriminative cut-offs, concurrent criterion validity, and test-retest reliability of the Oxford Vaccine Hesitancy Scale (OVHS): Chapter 4

Next, we report further details regarding the OVHS, including its discriminative performance, ideal cut-off choices associated with failure to be vaccinated against COVID-19, and test-retest reliability. These insights provide both important clinical cut-offs as well as methodological approaches that can be adopted when assessing other scales in the context of multiple independent datasets.

1.4.6 The Oxford Benchmark Scale for Rating Vaccine Technologies (OBSRVT): Chapter 5

Next, we report our development and validation of the Oxford Benchmark Scale for Rating Vaccine Technologies (OBSRVT). This instrument assesses the perceived acceptability of alternatives to needle and syringe-based vaccine delivery by benchmarking responses to needle and syringe-based vaccination.

1.4.7 Public attitudes towards ultrasound-mediated and intranasal vaccination: Chapter 6

We next report our findings comparing public preferences for ultrasound-based vaccination and intranasal vaccination. In addition to the OVHS and OBSRVT, we also assessed stated preferences and stated willingness to pay and performed two binary discrete choice experiments to better understand the relative valuation of these alternative vaccine delivery approaches.

1.4.8 Towards a Biotechnology De-Risking Toolkit: Chapter 7

Finally, the ways in which these approaches together represent a de-risking toolkit that can be adopted by other research teams are outlined. We highlight that an early focus on public attitudes and preferences can help de-risk novel biotechnologies and suggest that this could allow research groups to shift their focus from the *minimum viable product* to the *minimum desirable product*, ultimately lowering barriers to adoption and representing an early-warning system for technology viability while providing positive externalities.

1.5 Summary

A review of the literature suggests that pre-existing instruments do not address or define the impact of the myriad contributions of multiple domains to needle fear, vaccine hesitancy, and alternatives to needle and syringe-based vaccination. Validated instruments that explore these concepts and approaches are helpful for designing and engineering novel approaches to vaccine delivery and ensuring funder and public buy-in. The ONE scale, OVHS, and OBSRVT are platform-

agnostic instruments for de-risking vaccine delivery technologies. As shown in Chapter 6, participants—especially those with higher levels of vaccine hesitancy—preferred ultrasound-mediated vaccination over both needle and syringe and intranasal vaccination, underscoring the potential clinical utility of this technology. Beyond evaluating the potential benefits of future alternative approaches to vaccination, these methods and instruments may ultimately be broadly useful in developing behavioural strategies to decrease vaccine hesitancy in the UK, the US, and worldwide and assist in the successful adoption and deployment of alternative vaccine delivery technologies.

Chapter 2: The Oxford Needle Experience (ONE) Scale

2.1 Introduction

Vaccine hesitancy was identified by the World Health Organization as one of the top ten threats to public health even before the onset of the COVID-19 pandemic, (World Health Organization, 2019). Given the importance of behavioural considerations in modulating vaccine acceptance, changes in methods of vaccine delivery may decrease vaccine hesitancy, leading to significant global health benefits (Hettinga & Carlisle, 2020). With upwards of 10 billion vaccinations given by needle each year, better understanding the role of needles as potential contributors to vaccine hesitancy is of urgent public health importance, particularly as alternative approaches for vaccine delivery are undergoing development (Mathieu et al, 2021; McMurtry et al, 2015; Miller & Pisani, 1999).

Despite the central importance of understanding public attitudes to needle use for vaccination and other medical procedures, most previous scales designed to study needle attitudes, experiences, and fear were developed without fundamental qualitative research, in the absence of a multidimensional approach—often focusing exclusively on needle fear, anxiety, or even phobia—or in populations with suboptimal generalizability (Hako et al, 2022; Kose & Mandiracioglu, 2007; Olatunji et al, 2012; Öst et al, 1992; Ovanessian et al, 2019; Page et al, 1997a; van Overveld et al, 2011; Wenzel & Holt, 2003).

Validated multidimensional scales are a useful tool to study behavioural phenomena, and may provide insight with direct global health implications (Kantor & Kantor, 2021). Our goal, therefore, was to develop a feasible multidimensional scale that would capture the key domains contributing to concern regarding needle use in the context of vaccination or medical procedures. By better elucidating the underlying domains that may modulate vaccine hesitancy, and medical procedure hesitancy more broadly, these data may ultimately be used to better inform the development of evidence-based vaccination strategies to mitigate some of the perceived negative aspects of vaccine delivery. Moreover, a better understanding of the contribution of needles to vaccine hesitancy may also encourage the development of alternative vaccine-delivery technologies, ultimately resulting in more tailored and publicly acceptable approaches to vaccination and medical procedures in the future.

2.2 Methods

2.2.1 Domain and Item Generation

Our general framework and methodology is based on the existing literature regarding best practices for scale development (Boateng et al, 2018) and includes three sequential phases: item development (domain specification, item generation, and content validation), scale development (pre-testing, sampling, and factor analysis), and scale evaluation (reliability and validity assessment). This approach has been widely adopted in health services research and applied across diverse contexts including water insecurity (Young et al, 2019), treatment expectations (Alberts et al, 2020), and even vaccine hesitancy (Shapiro et al, 2018).

This methodology is also consistent with comprehensive guidance provided by Watkins (2022) on best practices for exploratory and confirmatory factor analysis in scale development when using Stata. Psychometric term definitions are included in the Appendix.

Domain generation to establish the domains of interest for the ONE scale was accomplished through an iterative deductive and inductive approach. After a literature review was performed, surveys were used to qualitatively develop a set of possible themes, and ultimately domains, that putatively contribute to needle attitudes, experiences, and expectations. Initial surveys were conducted iteratively using two separate 50-respondent samples recruited through Prolific Academic. The first was a convenience sample in the UK, and the second included a group of respondents who had previously answered that they had not received any vaccinations for COVID-19. The subject was introduced by stating that 'The aim of this study is to create a survey that can be used to understand attitudes towards needles and vaccine administration via needles. The reason we are asking you to complete the survey is that to develop valid surveys for future use it is important for researchers to determine whether the questions that we are asking are reasonable, relatable, and understandable. It is also important for us to look at the ways in which asking similar questions in slightly different ways may affect responses, and how the answers to these questions group together. This will enable us to develop meaningful survey instruments in the future that can be used to better assess public attitudes.' Open ended questions included items such as 'What are five words that come to mind when you think about needles?' In addition, respondents were presented with multiple

choice questions and asked to provide feedback on question quality, clarity, and saliency.

To assess the face and content validity of the scale, extensive qualitative measures were performed. The two 50-member online survey groups were asked, 'Did you feel like this set of multiple-choice questions was easy to understand? Reasonable?' to gain insights into the clarity, relevance, and interpretability of the items. This feedback mechanism was instrumental in refining the items to ensure they are aligned with the intended constructs and are easily understood. Furthermore, the iterative workshopping of items among the investigators, the Oxford Vaccine Group, the full BUBBL group at the Department of Engineering Science, and biology/ in vivo cluster meetings enriched the scale's content validity. Each item was reviewed and refined through this collective feedback process incorporating respondent input on clarity and relevance alongside iterative workshopping to ensure a comprehensive and representative measure of participants' perceptions and experiences regarding needles and injections.

Items were derived from three complementary sources. First, the literature review identified constructs assessed by existing needle scales, primarily fear, anxiety, and physiological responses to injection (Kleinknecht et al, 1999; Kose & Mandiracioglu, 2007; Öst et al, 1992; Page et al, 1997b) and identified possible untapped potential domains. For example, prior instruments did not focus on recalled negative experiences as a distinct contributor to needle attitudes, the potentially mitigating role of positive beliefs about needle utility, and broader attitudinal dimensions beyond fear or phobia that may be relevant in general (non-phobic)

populations. Second, open-ended responses from the two pilot samples generated candidate themes inductively, highlighting priorities and language that might not be apparent to investigators a priori, an approach designed to move beyond a reliance on investigator-derived a priori frameworks (Betsch et al, 2018; Frew et al, 2019). Third, iterative workshopping with the expert advisory groups refined both domain boundaries and item wording, ensuring clinical and public health relevance. Items addressing recalled negative experiences and perceived benefits of needles were specifically crafted to address dimensions absent from prior instruments, while items within the injection hesitancy and blood-related hesitancy domains were informed by constructs established in the existing literature but reworded to reflect the language and priorities identified by pilot study participants.

Participant item responses were recorded on a 5-point Likert scale ranging from strongly agree to strongly disagree. Items were reverse scored, as appropriate, to yield a scale where higher values represent greater levels of concern regarding needle use. This work was approved by the University of Oxford Medical Sciences Interdivisional Research Ethics Committee (approval reference R81585/ RE001). All subjects provided informed consent for participation and were permitted to withdraw from their anonymous surveys at any time. Qualitative analyses were performed using NVivo release 1.4 (QSR International, Burlington, Massachusetts, USA) and statistical analyses were performed using Stata for Mac version 16 (Stata Corporation, College Station, Texas, USA).

2.2.2 Data Collection

Two separate representative samples were used in the development of the ONE. A sample of 500 respondents in the United Kingdom (UK) was recruited through Prolific Academic (Oxford, UK), an academic survey panel site, and a draft scale was presented using the Qualtrics platform (Qualtrics Corp, Provo, Utah, USA). All respondents were rewarded with a small payment (<£2). This sample was used to develop the scale and perform exploratory factor analysis (EFA). A second demographically representative sample of respondents from the United States (US) was then solicited in a similar fashion using the same electronic survey recruitment tools. This sample was used for confirmatory factor analysis and reliability and validity testing. Results from the US and UK were evaluated separately to establish the robustness of the scale. Demographic information was self-reported by respondents and was linked with responses through an anonymous 25-digit alphanumeric key. Baseline information included age, self-reported sex, self-reported race and ethnicity, employment status, education history, household income, and vaccination status against COVID-19.

Prolific Academic's quota sampling functionality was used to roughly match the UK sample to Office of National Statistics population estimates and the US sample to 2015 US Census Bureau data. Marginal quotas were set for age (in 10-year bands), sex, and ethnicity to approximate population distributions. While this approach does not constitute true stratified random sampling, which would require random selection within pre-defined strata from a sampling frame, quota sampling provides a practical means of achieving demographic balance in online panel research and is widely used

in scale development and validation studies (Peer et al, 2017). Respondents who matched the required demographic profile were eligible to participate until quotas were filled, ensuring that the final sample approximated national demographic distributions on the specified demographic characteristics.

2.2.3 Domain and Item Reduction and Exploratory Factor Analysis

Bartlett's test of sphericity and the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was used to assess whether the correlation matrix was suitable for factor analysis (Kaiser, 1974; Lorenzo-Seva & Ferrando, 2021). The KMO statistic quantifies the proportion of variance among variables that may be common variance—that is, variance that might be attributable to underlying factors. It compares the sum of squared correlations to the sum of squared correlations plus squared partial correlations; when partial correlations are small relative to the original correlations, the KMO value approaches 1.0, indicating that correlations between pairs of variables can be explained by the other variables and thus that factor analysis is likely to yield distinct, meaningful, and reliable factors. The KMO is not a hypothesis test but rather a descriptive index: values above 0.90 are conventionally interpreted as 'marvellous,' values between 0.80 and 0.89 as 'meritorious,' and values between 0.70 and 0.79 as 'middling' (Kaiser, 1974). Values below 0.50 are 'unacceptable,' suggesting the data may not be appropriate for factor analysis. The term sampling adequacy thus refers not to sample size per se, but to whether the sample provides adequate shared variance for factor extraction.

To establish the number of domains to retain for the ONE scale, we calculated the minimum average partial correlation for the number of principal components, developed a scree plot, and performed parallel analysis for principal components (Dinno, 2009; Horn, 1965; Watkins, 2022). The minimum average partial correlation analysis successively 'partials' out shared variance among items, retaining the number of components at which the average squared partial correlation is minimized—the point at which the remaining correlations among items are as close to zero as possible (Watkins, 2022). The scree plot displays eigenvalues in descending order, with the number of factors to retain indicated by the point at which the curve inflects, or elbows. Factors above this inflection account for meaningful variance, while those below represent noise (Watkins, 2022). Parallel analysis compares *observed* eigenvalues against those generated from random data of the same dimensions; factors are retained only when the observed eigenvalue exceeds the corresponding random eigenvalue, indicating that the factor captures more variance than would be expected by chance (Dinno, 2009; Horn, 1965).

Item reduction followed a sequential, theory-driven process, consistent with best practices in scale development (Cappelleri et al, 2014). First, items were evaluated for domain relevance, defined as theoretical alignment with the general construct of needle attitudes. Items judged to be conceptually peripheral (such as items addressing general medical anxiety without specific reference to needles) were removed at this stage. Second, items with low polyserial correlations (item-total correlations below 0.30) were flagged for potential removal, since such items contribute little to the overall scale score. Third, polychoric inter-item correlations

were examined to identify redundancy; when pairs of items correlated above 0.80, the item with weaker theoretical justification or lower item-total correlation was removed to reduce respondent burden without sacrificing content coverage. Fourth, during exploratory factor analysis, items with communalities below 0.40 (indicating that less than 40% of item variance was explained by the extracted factors) or primary factor loadings below 0.40 were removed, as these items do not adequately represent the underlying latent constructs (Costello & Osborne, 2005). Finally, items with substantial cross-loadings (secondary loadings exceeding 0.30) were evaluated for theoretical coherence and removed a priori unless they reflected a unique clinically relevant construct. At every phase, the clinical cohesiveness of the domains and items were prioritized, particularly when results were borderline. This hierarchical approach prioritized theoretical coherence before statistical criteria, then applied statistical thresholds sequentially to produce a parsimonious scale with clear factor structure.

EFA with categorical variables was performed by using polychoric correlations to generate a factor matrix using iterated principal factor analysis followed by promax oblique factor rotation (Flora & Curran, 2004; Holgado–Tello et al, 2010).

Polychoric correlations were computed from the raw Likert-scale responses, and EFA was performed on the resulting polychoric correlation matrix rather than a traditional Pearson correlation matrix. This approach is essential when analysing ordinal data as Pearson correlations assume continuous, normally distributed variables and therefore may systematically underestimate the strength of association between ordinal items, leading to attenuated factor loadings and potentially incorrect factor structures (Flora & Curran, 2004; Holgado–Tello et al, 2010). Polychoric

correlations thus estimate the correlation between the latent continuous variables assumed to underlie the observed ordinal responses, providing more accurate, though computationally intensive, input for factor analysis.

Second, iterated principal factor analysis was employed as the extraction method. Unlike principal components analysis, which aims to explain total variance including unique variance, principal factor analysis focuses specifically on shared (common) variance among items, making it more appropriate when the goal is to identify the *latent* constructs underlying observed responses (Fabrigar et al, 1999).

Third, promax oblique rotation was selected rather than orthogonal rotation methods such as varimax. Oblique rotation explicitly permits factors to correlate, which is theoretically appropriate when—as with needle attitudes—the underlying constructs are anticipated to be related rather than independent. For instance, fear of needles and avoidance of injections are conceptually linked, and therefore forcing orthogonality would produce a distorted factor structure. Promax specifically provides a computationally efficient oblique rotation that begins with a varimax solution and then relaxes the orthogonality constraint, which may yield a more simple structure than direct oblique methods (Hendrickson & White, 1964). Our factor correlation matrix (Table 2.2) confirmed that the factors were indeed correlated, supporting our choice of oblique rotation.

2.2.4 Confirmatory Factor Analysis

The robustness of the findings was tested by performing confirmatory maximum likelihood factor analysis using structural equation modelling on the US

sample with the model developed in the UK sample with a Satorra-Bentler scaled test statistic. Multiple goodness of fit indices were analysed, including the root mean squared error of approximation (RMSEA), the standardized root mean squared residual (SRMR), and the comparative fit index (CFI) to assess model fit. Values of RMSEA≤0.08, SRMR≤0.08, and CFI≥0.90 suggest acceptable fit (Browne & Cudeck, 1992).

Confirmatory factor analysis was conducted within a structural equation modelling (SEM) framework because SEM provides a flexible approach to both specifying and testing hypothesised models. Unlike EFA, which identifies factor structures empirically, CFA requires the researcher to specify a priori which items load onto which factors, permitting a formal test of whether the hypothesized structure (in our case derived from the UK EFA) adequately reproduces the observed covariance matrix in an independent sample (the US sample). SEM software using Stata's sem command enables simultaneous estimation of all model parameters and provides standardized fit indices for model evaluation.

The Satorra-Bentler scaled chi-square statistic was employed since standard maximum likelihood chi-square tests assume multivariate normality, an assumption frequently violated with Likert-scale data, which are bounded, ordinal, and often skewed (Satorra & Bentler, 2010). When data deviate from normality, the standard chi-square is inflated, leading to excessive model rejection and the Satorra-Bentler correction adjusts the chi-square statistic by a scaling factor derived from the kurtosis of the observed variables, providing a more accurate test of model fit under non-normality. Robust standard errors are also computed, yielding more trustworthy

parameter estimates and confidence intervals, a correction widely recommended for CFA with ordinal variables (Flora & Curran, 2004).

Model fit was evaluated using multiple complementary indices, as no single index captures all aspects of fit (Jackson et al, 2009). The root mean square error of approximation (RMSEA) estimates the *discrepancy* between the hypothesised model and the population covariance matrix per degree of freedom, and rewards parsimony by penalizing model complexity. Typical interpretations suggests that RMSEA values ≤0.05 indicate close fit, values between 0.05 and 0.08 indicate reasonable fit, and values >0.10 indicate poor fit (Browne & Cudeck, 1992). The standardized root mean square residual (SRMR) represents the average standardized difference between observed and predicted correlations; values ≤0.08 are generally considered acceptable, while values ≤0.05 indicate excellent fit (Hu & Bentler, 1999). The comparative fit index (CFI) compares the hypothesized model to a baseline model in which all variables are uncorrelated, There, values ≥0.95 suggest excellent fit, while values between 0.90 and 0.95 suggest acceptable fit (Hu & Bentler, 1999). We adopted the two-index strategy recommended by Hu and Bentler (1999), requiring acceptable values on both an absolute fit index (RMSEA or SRMR) and an incremental fit index (CFI) to establish adequate model fit.

2.2.5 Reliability and Validity

Reliability as internal consistency was assessed by examining McDonald's omega, a more robust alternative to Cronbach's alpha, for each subscale domain (factor) and for the overall scale, in both the UK and US samples (Streiner, 2003). The

subtotal of items in each of the domains in the scale represent a subscale that addresses that domain in particular, and the CFA assesses the way these subscales (as factors) meaningfully generalize to the US population.

Convergent validity was evaluated by examining the correlation between the ONE scale and the four-item OCEAN fears factor score (Freeman et al, 2022). Predictive validity was assessed by examining the association between the ONE scale score and self-reported vaccination status against COVID-19. Logistic regression odds ratios of association were calculated with vaccination status (defined as having received a minimum of one dose of any vaccine for COVID-19) coded as a dichotomous variable and the ONE scale treated as a continuous predictor variable. Discriminant validity was assessed by examining the logistic regression odds ratios of association between the ONE score and agreement with a non-specific statement ('it is important to control what enters one's body') in the UK and US samples. A p-value of 0.05 was considered significant.

2.3 Results

2.3.1 Respondent Characteristics

In the UK sample, of the 506 respondents who started the survey, 500 subjects returned completed surveys, yielding a 98.8% completion rate in the survey panel. The median (interquartile range [IQR]) age was 45 (32, 59), and respondents ranged in age from 18 to 90. 50.4% of respondents identified as female, 48.2% as male, and 1% as non-binary. The median (IQR) self-reported household income was £39,000

(25,000, 58,000). A total of 47.6% of respondents were currently working full-time and 14.2% identified as non-white.

In the US sample, of the 509 respondents who started the survey, 500 subjects returned completed surveys, yielding a 98.2% completion rate in the survey panel. The median (IQR) age was 45 (31, 60), and respondents ranged in age from 18 to 84. 49.6% of respondents identified as female, 47.6% as male, and 1.6% as non-binary. The median (IQR) self-reported household income was $53,500 (31,000, 89,999). A total of 45.5% of respondents were currently working full-time and 27.8% identified as non-white.

2.3.2 Domain and Item Reduction and Exploratory Factor Analysis

A total of 16 themes were identified as potentially contributing to public attitudes to needles after literature review and qualitative analysis of initial pilot survey data. The themes identified through a literature review included fear, history of negative experiences, intrusive thoughts, risk, avoidance, needle appearance, fainting, anxiety, pain, blood, and trust. The initial pilot survey feedback added several additional themes including general dislike, concerns regarding personal control, fears regarding maintaining bodily integrity, positive aspects of needle use, and the training/expertise/skill of the injector. A total of 107 items for potential scale inclusion were then developed for testing; the intentionally large domain and item pool was designed to pilot similar iterations of items in a pilot survey format. A total of 88 items were sequentially deleted due to low domain relevance (12), low polychoric or

polyserial correlations (22), unifactorial outcomes (3), redundancy (33), or low communality (18).

Factor loadings from the 19 retained items after EFA for categorical variables with promax oblique rotation are presented in Table 2.1. Bartlett's test of sphericity ($p<0.0001$) suggested that the data are appropriate for EFA; this was confirmed by the KMO measure of sampling adequacy, which was excellent, with a value of 0.92 for the overall model and with values ranging from 0.68 to 0.97 for each of the individual items. The scree plot, minimum average partial correlation analysis, and the eigenvalue > 1 criterion converged on a 4-factor solution, representing four latent constructs underlying the item responses. This solution reflected a simple structure, meaning that each item loaded strongly (≥0.50) onto a single factor with minimal cross-loadings onto the remaining factors (Table 2.1), a pattern suggesting that items cluster into distinct, interpretable domains rather than overlapping across multiple constructs (Garrido et al, 2011; Ledesma et al, 2015). Parallel analysis suggested retaining three factors, though this result was considered borderline given the adjusted eigenvalue of 0.97 for the fourth factor, just below the cut-off of one (Dinno, 2009).

Domain	Item	Factor I	Factor II	Factor III	Factor IV
Injection	I am afraid of needles.	**0.9449**	0.0468	-0.0622	-0.0951
	I am afraid of being vaccinated because of the needle.	**0.9808**	-0.1095	-0.0307	0.0119
	I have experienced severe panic when thinking about or undergoing an injection.	**0.7896**	0.0227	0.1669	0.0064

	When I know I will need to be injected with a needle, I worry about it ahead of time.	**0.9088**	-0.0146	0.0406	-0.0914
	I have delayed getting a vaccine specifically because of the needle.	**0.7954**	-0.0051	0.0474	0.1659
	I am bothered by the appearance of a needle.	**0.7300**	0.2277	-0.0493	-0.0410
	I feel anxious at the thought of being injected with a needle.	**0.8713**	0.1421	-0.0257	-0.0505
	I am worried that I will feel unwell or faint if I am injected with a needle.	**0.5640**	0.2972	0.0736	-0.0564
	Needles cause significant pain.	**0.6236**	-0.1182	0.0984	0.1585
	I dislike that a needle pierces my skin and enters my body directly.	**0.7196**	-0.0200	0.0426	0.0897
Blood	I am afraid of seeing blood.	0.2112	**0.6760**	-0.0008	0.0197
	I avoid watching surgical procedures on television or on the internet.	0.1448	**0.5628**	-0.0812	-0.0174
	I feel lightheaded when I see blood.	-0.0415	**0.9135**	0.0460	0.0736
History	I had a bad experience with needles in the past.	0.0861	0.0572	**0.8677**	0.0616
	I have a strong unpleasant memory of being injected with a needle.	0.3179	-0.0119	**0.7238**	-0.0835
	I remember a nurse or doctor making a mistake when injecting me with a needle.	-0.1582	0.0021	**0.8672**	-0.0221
Benefits	I think needles are useful.	0.1967	-0.0099	-0.0248	**0.5859**
	My feelings regarding needles do not matter since needles are so useful.	-0.0694	0.0645	0.0150	**0.7389**
	People should try to overcome their fear of needles because of the benefits of using them.	-0.0999	0.0274	-0.0146	**0.8757**

Table 2.1 Factor loadings for the ONE exploratory factor analysis (EFA).

Factors correlated clinically with the domains of 1) injection hesitancy (including fear and anxiety related to injection); 2) blood-related hesitancy; 3) recalled negative experiences; and 4) perceived benefits of needles.

Factor intercorrelations are included in Table 2.2, and the final ONE scale and scoring key are presented in Table 2.3.

	Injection	**Blood**	**History**	**Benefits**
Injection	1			
Blood	0.59	1		
History	0.50	0.20	1	
Benefits	0.36	0.20	0.15	1

Table 2.2 Factor intercorrelations in the UK EFA.

Number	Item	Domain
1	I am afraid of needles.	Injection
2	I am afraid of being vaccinated because of the needle.	Injection
3	I have experienced severe panic when thinking about or undergoing an injection.	Injection
4	When I know I will need to be injected with a needle, I worry about it ahead of time.	Injection
5	I have delayed getting a vaccine specifically because of the needle.	Injection
6	I am bothered by the appearance of a needle.	Injection
7	I feel anxious at the thought of being injected with a needle.	Injection
8	I am worried that I will feel unwell or faint if I am injected with a needle.	Injection
9	Needles cause significant pain.	Injection
10	I dislike that a needle pierces my skin and enters my body directly.	Injection
11	I am afraid of seeing blood.	Blood
12	I avoid watching surgical procedures on television or on the internet.	Blood
13	I feel lightheaded when I see blood.	Blood
14	I had a bad experience with needles in the past.	History
15	I have a strong unpleasant memory of being injected with a needle.	History
16	I remember a nurse or doctor making a mistake when injecting me with a needle.	History
17	I think needles are useful.	Benefits
18	My feelings regarding needles do not matter since needles are so useful.	Benefits
19	People should try to overcome their fear of needles because of the benefits of using them.	Benefits

Table 2.3 The ONE Scale. All answer choices are rated using a 5-point Likert scale (strongly disagree to strongly agree). For questions 1-16, strongly agree is scored as 5, while for questions 17-19 strongly agree is scored as one (reverse scoring). Higher values reflect more hesitancy regarding needle use. The total possible ONE score ranges from 19 to 95.

2.3.3 Confirmatory Factor Analysis

Confirmatory factor analysis was performed on the US sample using the model developed in the UK sample. The maximum likelihood factor analysis model created using structural equation modelling demonstrated acceptable goodness of fit characteristics, with RMSEA=0.070 (90% confidence interval 0.064, 0.077), SRMR=0.053, and CFI=0.925 using the four-factor model with a total of 19 items.

The median (interquartile range [IQR]) total ONE score was 37 (29-48) in the UK population and 39 (29-54) in the US population (p=0.054). Further descriptive statistics for each subscale are included in Table 2.4.

	UK Sample Median (IQR)	**US Sample Median (IQR)**	**p-value**
Injection	17 (13-26)	19 (13-29)	0.07
Blood	7 (4-10)	7 (4-9)	0.80
History	6 (4-10)	8 (4-12)	<0.001
Benefits	5 (4-6)	5 (4-7)	0.32
ONE Scale Overall	37 (29-48)	39 (29-54)	0.054

Table 2.4 Median and interquartile ranges for ONE sub-scales and the overall scale.

P-values are for the two-sample Wilcoxon rank-sum (Mann-Whitney) test.

2.3.4 Reliability and Validity

Reliability as internal consistency was excellent: McDonald's omega for the overall ONE scale was 0.92 in the UK sample and 0.93 in the US sample. Omega values for each subscale are included in Table 2.5.

Subscale	UK Sample	US Sample
Injection	0.94	0.93
Blood	0.79	0.79
History	0.85	0.84
Benefits	0.71	0.70
Overall	**0.92**	**0.93**

Table 2.5 Subscale and overall reliability, as measured by McDonald's Omega.

Convergent validity with the OCEAN needle fear score supported our validity argument, showing a strong correlation between the scales (r=0.83). Predictive validity with the single-question vaccination status question demonstrated a significant association between vaccination status and ONE score, with a logistic regression odds ratio of association (95% CI) of 0.98 (0.96, 1.0), p=0.014 for the UK sample and 0.97 (0.96, 0.98), p<0.0001 for the US population. Given that the standard deviation of the ONE score in the UK and US samples is 15.2 and 16.5, respectively, this translates into a decrease in the odds of being vaccinated of 30.4% in the UK and 49.5% in the US for each single standard deviation increase in ONE score.

Discriminant validity with a question regarding the importance of controlling what enters the body confirmed that the ONE score does not predict this unrelated outcome, with logistic regression odds ratios of association of 1.01 (1.00, 1.02), p=0.162 in the UK population and 1.00 (0.99, 1.01), p=0.996 in the US sample.

2.4 Discussion

We have shown here that the ONE scale is a reliable multidimensional instrument and that there is significant preliminary evidence to support the scale's validity for assessing public attitudes to needle use in the context of medical procedures such as vaccination. Needle concerns are not limited to those undergoing vaccination, and have been implicated in hesitancy regarding a range of procedures from venepuncture (Deacon & Abramowitz, 2006) to local anaesthetic injection in dermatologic surgery (Patakfalvi & Benohanian, 2014), blood donation (Zucoloto et al, 2019), dental procedures (Berge et al, 2016), and dialysis (Radisic et al, 2022). Therefore, a better appreciation of the myriad contributors to the needle exposure experience may be broadly valuable.

In evaluating descriptive statistics for each of the subscales, only the median scores for a recalled history of negative experiences differed significantly between the UK and US populations, with an absolute 33% increase reported in the US group ($p<0.0001$). As such recall has been previously implicated as a potential driver of needle fear—where a feedback loop of recalled negative experiences leads to avoidance, furthering fear—this may have important implications for needle acceptance (McMurtry et al, 2015). Moreover, the magnitude of the association between ONE score and COVID-19 vaccination status—the odds of being unvaccinated increase by almost 50% for each standard deviation increase in ONE score in the US population—not only highlights the predictive validity of the scale but also underscores its utility and potential value in designing public health interventions and

the potential role of needle hesitancy as a contributor to vaccine hesitancy (Freeman et al, 2021).

2.4.1 Prior instruments

Several prior instruments have been developed to assess needle attitudes. The 18-item Injection Phobia Scale-Anxiety (IPS-Anx) was developed in 1992 on a sample of 40 subjects with known needle phobia and examines fear and anxiety exclusively (Öst et al, 1992). Two decades after its initial development, the factorial structure of the scale was assessed, suggesting a two-factor structure, encompassing contact (direct injection) and distal (watching or discussing injection) fears (Olatunji et al, 2010). The 50-item Medical Fear Survey examines an array of fears relating to medical procedures, including injections as well as blood, physical examination, and mutilation;(Kleinknecht et al, 1999) an abbreviated version of this survey has also been developed (Olatunji et al, 2012). The blood/injection fear scale was developed on a Turkish population of patients attending a hospital for examination or blood donation and is a 20-question scale that includes two domains, injection and blood (Kose & Mandiracioglu, 2007). The Blood-Injection Symptom Scale assesses domains of faintness, anxiety, and tension, and relied on a population of Australian university students for its development (Page et al, 1997a). The 40-item Multidimensional Blood/ Injury Phobia Inventory (MBPI) was developed on a US university student population to assess phobias, and is aimed at assessing pathological reactions, rather than understanding the range of experiences of the general population exposed to needles (van Overveld et al, 2011; Wenzel & Holt, 2003). Blood-injection-injury (BII)

was also assessed as part of a 10-item subscale of the Specific Phobia Questionnaire, a lengthy scale addressing phobias ranging from fear of animals to the natural environment; this scale was developed on a cohort of patients with known anxiety disorders and a group of Canadian university students (Ovanessian et al, 2019). Finally, the multidimensional fear of injection scale was developed on 419 Japanese university students to evaluate fear of injections exclusively, encompassing domains of direct fear, indirect fear, physiological response, and avoidance (Hako et al, 2022).

2.4.2 Study strengths

Unlike these scales, we used a qualitative methodology to ensure that our domain- and item-generation would include the diversity of priorities of a wide range of respondents to the experience of undergoing injection with a needle—a methodology designed to capture the broadest possible themes that may not be obvious to investigators a priori (Alberts et al, 2020; Boateng et al, 2018; Morgado et al, 2017; Rowan & Wulff, 2007). This is important, as prior studies have suggested a disconnect between clinicians' perception of patient or public priorities and their actual priorities (Boivin et al, 2014; Sathanapally et al, 2020). The ONE scale is also the only instrument to include domains that assess recalled negative experiences as well as positive beliefs regarding the utility of needles, which may add significantly to researchers' understanding of patient and public priorities.

Additional strengths of our study include the iterative approach and the use of representative populations in our exploratory and confirmatory factor analysis; many prior studies relied on less diverse populations of young, well-educated students—or,

conversely, on those being treated for anxiety disorders—and therefore the generalizability of their scales to the wider population is unknown. Furthermore, some prior scales aimed to evaluate pathological responses to needle injection—needle phobia. While such work is highly valuable, our goal was to develop a scale that could be broadly adopted for designing public health interventions and alternative vaccine and medical therapeutic delivery technologies for the public.

We also relied on statistical analyses appropriate for ordinal data, including a polychoric correlation-matrix based EFA, to provide conservative estimates given concerns regarding the assumed normality of Likert-scale data (DeWees et al, 2020; Wu & Leung, 2017). This represents a departure from earlier studies that acknowledged the skewness inherent in needle fear assessments—the majority of individuals have only mild symptoms while a small proportion have extreme reactions—but used statistics more appropriate for analysing normally-distributed data with continuous endpoints (Wenzel & Holt, 2003).

The samples from the UK and US were designed to be demographically representative. Prolific Academic creates representative samples in the UK and US populations by stratifying by age (across five 9-year age brackets), sex (male or female), and ethnicity (using the five simplified categories recommended by the UK Office of National Statistics: White, Mixed, Asian, Black and Other). Cross stratifying by age (5 brackets), sex (2 groups), and ethnicity (5 groups) yields 50 subgroups. The number of respondents in each subgroup is designed to match census data from the US Census Bureau (2015) or the UK Office of National Statistics (2011), as appropriate, and allocations to each subgroup are made proportionately. Subjects are recruited

from the survey panel site, whose current database includes over 130,000 identity-confirmed participants, and subjects are offered to participate if the particular subgroup to which they belong (for example, an Asian male between 18 and 27) has not yet been filled. We acknowledge that a panel-based sampling method has biases, though the large sample size and broad range of respondent characteristics suggests that this approach yields insights that are broadly applicable to the general populations of the UK and US.

2.4.3 Study limitations

Our work has several important limitations. First, we relied on a survey panel design to allow for a demographically representative population from both the UK and US, but there may be important differences between the general population and the subset willing to volunteer to contribute to survey panels, and response rates cannot be calculated with this approach. While completion rates were high (98.8% UK, 98.2% US), traditional population-based response rates cannot be calculated when using an open-enrolment panel design, as no defined sampling frame is contacted directly and therefore the true denominator of eligible individuals who declined participation is unknown. Moreover, all research that relies on self-report introduces the risk that bias, and particularly response bias and social desirability bias, may affect stated preferences (Podsakoff et al, 2003). Nevertheless, the robustness of our findings, with total and subscale scores largely similar across both the UK and US populations, supports the generalizability of the scale across these populations. Second, one domain—injection hesitancy—accounts for 10 of the 19 items in the scale, while the

remaining domains each include three items. The unequal number of items across domains, however, was purposely designed to capture the presence of a primary domain of interest—injection hesitancy—while also acknowledging and evaluating the contribution of other domains. Indeed, some have suggested that a unitary construct underlies responses to needle injection (Wenzel & Holt, 2003), and it was our goal to expand on this central domain rather than attempt to over-extract factors that otherwise correlate closely. Third, additional validation testing, as well as test-retest reliability and cross-cultural validation studies, should be considered in the future to better understand the limits of the ONE scale's generalizability, particularly given the potential for cultural and spiritual beliefs to affect these results. Fourth, an additional limitation of our manuscript is the absence of invariance testing and multi-group comparisons. We did not assess whether the ONE scale's item scores are comparable between subgroups based on demographic variables, such as country of origin, age, and socioeconomic status, or determine whether the scale operates equivalently across different groups regarding the perceived item content and the dimensionality of underlying constructs. Future research including such analyses may aid in establishing the scale's robustness across diverse clinical and research settings. Finally, an added limitation of our methodological approach was the lack of formal cognitive interviews performed to establish face and content validity. While we acknowledge that our approach does not provide an absolute guarantee of item interpretation consistency, the thorough and iterative review and feedback process bolsters the face and content validity of the ONE scale. Though the pilot survey data and feedback from experts was instrumental in developing the scale, future studies including cognitive

interviews may be helpful to further support arguments for the validity of the ONE scale for assessing needle attitudes.

2.4.4 Conclusions

The ONE scale addresses four separate domains that contribute to the public's experience with needles: injection hesitancy, blood-related hesitancy, recalled negative experiences, and perception of the benefits associated with needle use. Distinguishing between these domains and appreciating their independence and interdependence may provide helpful insights in designing public health interventions. Indeed, as with any multidimensional scale, when interpreting the total ONE score it is important to appreciate that different domains (subscales) contribute in varying amounts to the scoring and may offset each other. Injection hesitancy accounts for the lion's share of the variance, and includes 10 of the 19 total score items, but this can be offset by scores on the other domains—that is, if someone believes strongly in the benefits of needles, this could modulate their total score. While the correlations between the injection subscale and blood-related hesitancy subscale is fairly strong (0.59), some of the other subscales correlate poorly—for example, the correlation between the perceived benefits of needle use and other subscales ranges from 0.15 to 0.36. On a practical level, this means that while subscale data add important granular detail, they may also substantially offset each other, potentially diluting the actionable utility of the global score in understanding contributors to needle attitudes. Therefore, as with any multidimensional scale, the overall score provides useful summary information but for deeper insights attention must be directed at the

subscale scoring level. The ONE scale represents a new parsimonious instrument that may provide important actionable insights to promote global public health, and further research and testing may be warranted.

Chapter 3: Development and initial validation of the Oxford Vaccine Hesitancy Scale (OVHS)

3.1 Introduction

Vaccine hesitancy is a significant international public health priority, and better appreciating and predicting its aetiology, associations, and putative mitigating factors has taken on particular urgency in recent years (Cooper & Wiysonge, 2023; Eskola et al, 2015; Salmon et al, 2015b). Indeed, while concerns were raised globally regarding increasing levels of vaccine refusal for preventable childhood illnesses even prior to the COVID-19 pandemic (Dubé & MacDonald, 2016; The Lancet Child & Adolescent Health, 2019), the ensuing pushback against mandated, or even recommended, vaccination schedules (Carpiano et al, 2023) has only increased since vaccine hesitancy was labelled a top-10 threat to public health in 2019 (World Health Organization, 2019), as even primary care physicians admit a significant rate of mistrust in the safety and effectiveness of vaccines (Callaghan et al, 2022) while increasingly ubiquitous social media-heavy diets appear to further exacerbate hesitancy as well (Chadwick et al, 2021).

Several validated scales have been developed previously for evaluating vaccine hesitancy, and a scoping review has evaluated the past decade of scale development in this field (Oduwole et al, 2022). The Parent Attitudes about Childhood Vaccines (PACV) scale, the most highly cited scale in evaluating vaccine hesitancy (Bussink-Voorend et al, 2022), was developed to identify parents who are vaccine hesitant (Opel et al, 2011a; Opel et al, 2011b). While the PACV has important strengths, including its parsimony and multidimensional approach, it was developed exclusively

for childhood vaccinations and does not capture the rich multidimensionality of contemporary vaccine hesitancy across other groups.

The SAGE Vaccine Hesitancy Working Group survey includes three separate sets of items: core questions, Likert-scale questions, and open-ended questions (Larson et al, 2015). While these questions address a range of components that align well with the original SAGE definition of vaccine hesitancy, the factor structure of these questions was not explored and psychometric testing was not performed.

The Vaccine Confidence Scale (VCS) includes three domains (benefits, harms, and trust) over eight items, and the authors found that VCS scores were associated with vaccine refusal, though the small number of total items, particularly with three domains, means that it may not capture many of the complex and dynamic contributors to contemporary vaccine hesitancy (Gilkey et al, 2014; Gilkey et al, 2016).

The Vaccine Acceptance Instrument (VAI), a 20-item scale, examines five domains: safety, effectiveness, acceptance of scheduling, positive beliefs in vaccines, and the perceived legitimacy of those mandating vaccinations. Yet the small sample used to develop the scale (246 respondents) coupled with its psychometric limitations (principal components analysis suggested a three-factor solution which the authors rejected in lieu of their a priori theoretically-driven factor structure) means that its generalizability and robustness may be less than ideal.

The Vaccine Confidence Index (VCI) was developed specifically to assess parental trust in the US vaccination system (Frew et al, 2019), though its reliance on an a priori-designated 5-factor structure using only eight items suggests the need for multiple factors to load with a single item, likely yielding unstable results. The Vaccine

Attitudes Examination (VAX) scale is a 12-item four-domain instrument that addresses mistrust of vaccine benefits, worries regarding future unforeseen effects, concerns about commercial profiteering, and a preference for natural immunity (Martin & Petrie, 2017b). However, the use of university students and a non-representative convenience sample for development, and the small number (n=14) of participants included in the domain- and item-generation focus groups, may impair the generalizability of this scale.

The Vaccine Hesitancy Scale (VHS) was developed by a SAGE working group and included nine items over two domains: lack of confidence and risks (Shapiro et al, 2018). Yet in examining only these two interrelated concepts, and focusing exclusively on parental attitudes to childhood vaccination, this scale may not generalize well to other contemporary vaccination scenarios.

The 5C scale was developed to assess not only confidence, complacency, and convenience but calculation and collective responsibility as well; the authors theorized that if they could evaluate the 'psychological antecedents to vaccination'(Brewer et al, 2017) this would allow them to better understand the complex phenomenon of vaccine hesitancy (Betsch et al, 2018). Yet the reliance on an a priori-derived theoretical framework to explain the behavioural antecedents of vaccine hesitancy may limit the ways in which it captures the full range of domains contributing to hesitancy in the real world.

Finally, The Multidimensional Vaccine Hesitancy Scale (MVHS), a 32-item 8-domain scale developed on a survey-panel population, was created in an attempt to broaden the number of domains included in existing scales (Howard, 2022). However,

it suffers from important limitations including the use of non-representative samples across varying populations, the use of statistics appropriate for continuous data when assessing Likert-scale data, the lack of any trust- or conspiracy-related items, and significant redundancy which results in several items that are 'near synonyms' (Howard, 2022).

Given the limitations of pre-existing scales, we aimed specifically to practically, feasibly, and theoretically improve upon existing instruments by developing a scale supported by the literature on vaccine hesitancy but bolstered by the inclusion of expert opinion and qualitative input from populations of both the general population and the subpopulation of vaccine hesitant individuals. This approach allows us to better capture the range of experiences and expectations associated with the contemporary vaccine hesitancy landscape. By considering a wide array of domains and items and then rigorously applying a robust iterative statistical methodology to domain and item reduction, exploratory, and confirmatory factor analysis, this approach yields a feasible multidimensional scale with significant support for our validity argument.

Opting to forego vaccination may be the result of rational decision making, an emotional response, mistrust, or a combination of factors. The very heterogeneity of the umbrella concept of vaccine hesitancy—as evidenced by those across the ideological spectrum that consider themselves hesitant (Jiang et al, 2021; Stoeckel et al, 2022)—means that a multidimensional approach to understanding this phenomenon is crucial for public health professionals, behavioural scientists, and vaccine researchers. Moreover, quantifying vaccine hesitancy in an accurate, feasible,

and reproducible fashion may help better tailor future alternative vaccine delivery approaches and technologies to improve the likelihood of vaccine acceptance (Hettinga & Carlisle, 2020). We therefore aimed to develop a psychometrically sound, feasible, and multidimensional scale to assess the broad concept of vaccine hesitancy based on both the existing literature and qualitative research.

3.2 Methods

3.2.1 Data Collection and Population

As part of the pre-development process, a series of surveys including both open- and closed-ended questions were performed via the academic survey site Prolific Academic (Oxford, UK). The first cohort included a convenience sample of 50 respondents, while the second included a separate set of 50 respondents who reported that they had never received any vaccinations against COVID-19.

OVHS development was then performed using two distinct assessment cohorts. First, a demographically representative (based on UK Office of National Statistics matched stratification on age, sex, and ethnicity) sample of 500 participants from the UK was enlisted using Prolific Academic. The initial iteration of the scale was administered to these participants via the Qualtrics platform (Provo, Utah, USA). Each participant received a nominal payment of less than £4 for their contribution. The data collected from this UK sample was used for exploratory factor analysis (EFA) to refine the scale's structure and content. Psychometric term definitions are included in the Appendix.

Subsequently, a second sample consisting of demographically representative (based on 2015 US national census bureau data) US-based respondents was

conducted utilizing an identical electronic survey recruitment methodology. This second phase of data collection was conducted to perform a confirmatory factor analysis (CFA), to ensure the model's robustness across populations, and to assess the scale's reliability and validity. We elected to perform the CFA on a US sample rather than on a second UK sample to allow us to a) assess the robustness of the EFA-derived model on a separate population with a different cultural context; b) highlight the generalizability of our model; and c) efficiently leverage the use of a second population that could both be used to assess model goodness of fit while concomitantly highlighting its generalizability and robustness and avoiding the need for a third sample, consistent with best practices in scale development and our previous work in this area (Boateng et al, 2018).

Each participant's demographic details were self-declared and coupled with their responses using a unique 25-character alphanumeric identifier to maintain anonymity. The demographic data gathered encompassed a range of attributes, including age, self-identified gender, self-identified race and ethnicity, employment status, educational background, household income, and COVID-19 vaccination status.

Our overall methodological approach is similar to our earlier scale development manuscripts and the work outlined in the preceding chapter (Kantor & Kantor, 2021; Kantor et al, 2023). This work was approved by the University of Oxford Medical Sciences Interdivisional Research Ethics Committee (approval reference R81585/ RE001). All subjects provided informed consent for participation and were permitted to withdraw from their anonymous surveys at any time. Qualitative analyses were performed using NVivo release 1.4 (QSR International, Burlington,

Massachusetts, USA) and statistical analyses were performed using Stata for Mac version 16 (Stata Corporation, College Station, Texas, USA).

3.2.2 Domain and Item Generation

We approached domain and item generation through an iterative deductive and inductive approach. After a thorough literature review to generate a preliminary domain set, the two 50-member online surveys were used to broaden the range of domains for potential inclusion in the OVHS. Participants also provided feedback regarding sets of test items to gauge their interpretability and were offered the opportunity to note whether any items were difficult to understand, redundant, or should be reworded. Qualitative data from open- and closed-ended questions, along with feedback from a group of experts from the fields of medicine, public health, vaccinology, social science, and engineering science (obtained via workshop group discussions in both the Oxford Vaccine Group and the biology/ in vivo cluster meetings of the University of Oxford Department of Engineering Science), contributed to domain development and inclusion to better support face and content validity, an approach consistent with our previous work in this area and established scale development methodology (Boateng et al, 2018; Kantor & Kantor, 2021; Kantor et al, 2023; Watkins, 2022).

Preliminary responses were recorded using 5-point Likert scale responses and open-ended (free text) responses, as appropriate. Open-ended questions included items such as listing five words that come to mind when thinking about vaccines, as well as when thinking about other key stakeholders, such as the people, governments,

and organizations developing vaccines. These questionnaires also explored the degree to which respondents thought of vaccines as protecting public health and whether they believed vaccines are effective. For the representative samples, responses were recorded on a 5-point Likert scale with options ranging from strongly disagree to strongly agree. Items were reverse scored, as appropriate, ultimately yielding a scale where higher values represent greater levels of vaccine hesitancy.

3.2.3 Domain and Item Reduction and Exploratory Factor Analysis

Bartlett's test of sphericity and the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy were performed to ensure that the data were amenable for factor analysis (Lorenzo-Seva & Ferrando, 2021). To establish the ideal number of domains to retain for inclusion in the OVHS we relied on the minimum average partial correlation for the number of principal components, developing a scree plot, and performing parallel analysis for principal components (Dinno, 2009; Horn, 1965; Watkins, 2022). Additionally, item reduction was performed by assessing domain relevance, examining polychoric (item-item) and polyserial (item-test) correlations, culling redundant items, and removing those with low communality or low factor loadings (Cappelleri et al, 2014). To ensure that our approach yielded statistically meaningful results with non-continuous data, EFA with categorical variables was performed using polychoric correlations to generate a factor matrix using iterated principal factor analysis followed by promax oblique factor rotation (Flora & Curran, 2004; Holgado–Tello et al, 2010).

3.2.4 Confirmatory Factor Analysis

Confirmatory maximum likelihood factor analysis using structural equation modelling with a Satorra-Bentler scaled test statistic was performed on the US sample with the model developed in the UK sample to establish the robustness of our model. Goodness of fit indices analysed included the root mean squared error of approximation (RMSEA), the standardized root mean squared residual (SRMR), and the comparative fit index (CFI). Values of RMSEA≤0.08, SRMR≤0.08, and CFI≥0.90 suggest acceptable fit, while values of RMSEA≤0.05, SRMR≤0.06, and CFI≥0.95 suggest excellent fit (Browne & Cudeck, 1992).

3.2.5 Scale Reliability and Validity

Reliability as internal consistency was assessed using McDonald's omega, a more robust alternative to Cronbach's alpha that does not assume equal factor loadings. This was performed for the overall OVHS as well as each subscale (Streiner, 2003).

Predictive validity was assessed by evaluating the association between OVHS score and self-reported vaccination status against COVID-19. Logistic regression odds ratios of association were assessed per standard deviation change in OVHS score (the predictor variable) with vaccination status against COVID-19 (with positive vaccination status defined as a self-report of having received at least one vaccination) serving as the dichotomous outcome of interest.

Convergent validity was assessed by evaluating the correlation between the OVHS and the 5C psychological antecedents to vaccination scale (short version), a

validated 5-item scale assessing vaccine hesitancy (Betsch et al, 2018). Discriminant validity was assessed by examining the logistic regression odds ratios of association per standard deviation change in the OVHS score predicting agreement with a statement regarding preference for performing outdoor activities. For all analyses, a p-value of 0.05 was considered significant.

3.3 Results

3.3.1 Respondent Characteristics of Representative Samples

3.3.1.1 UK EFA Sample

For the EFA sample of UK respondents, of the 537 participants who initiated the survey link, 15 declined the survey, 11 timed out (did not complete), and seven respondents were rejected for failing attention checks. This yielded 504 complete acceptable responses, a 93.9% completion rate. The median (interquartile range [IQR]) age was 45 (32, 60), with respondents ranging in age from 19 to 81. Forty-nine percent of respondents identified as male, and the median (IQR) household income was £40,000 (25,000, 60,000). A total of 14.9% of the participants identified as non-white, and 47.9% reported working full-time.

3.3.1.2 US CFA Sample

For the CFA sample of US respondents, of the 513 participants that initiated the survey link, 10 declined the survey, two timed out (did not complete), and one respondent was rejected for failing attention checks. This yielded 500 complete responses, a 97.4% completion rate. The median (interquartile range [IQR]) age was 46 (31, 60), with respondents ranging in age from 18 to 82. Fifty percent of

respondents identified as male, and the median (IQR) household income was $68,000 (35,500, 110,000). A total of 21.4% of the participants identified as non-white, and 51.6% reported working full-time.

3.3.2 Domain and Item Reduction and Exploratory Factor Analysis

We identified 26 potential domains: perceived intent of vaccine stakeholders, intentions and willingness, cognition, vaccination behaviour, decision-making and indecisiveness, perception of vaccine benefits, social norms, trust in healthcare professionals, risk assessment, collective responsibility, concerns about pain, perceived profit motives, misinformation, development haste, risks and side effects, inequity, natural immunity preference, personal freedom and control, concern regarding unnecessary vaccinations, vaccination process concerns, past negative experiences, inner conflict, opposition to vaccines, needle fear, complacency, and convenience. Of these, 23 were based on the preexisting literature and three (vaccination process concerns, past negative experiences, and inner conflict) emerged inductively from directed content analysis of the open-ended qualitative survey responses. We considered a total of 137 items for inclusion in the OVHS after expert consultation, literature review, and initial open- and closed-ended questionnaire performance. These included positively oriented declarative items such as ‘The benefits of vaccines outweigh their risks’ and ‘Vaccines save lives’, negatively-oriented items such as ‘Getting vaccinated is painful’, and items that address anti-vaxevangelism such as ‘I would discourage my friends or family from getting vaccinated’. A total of 124 items were culled due to redundancy (58), low domain

relevance (15), low polychoric or polyserial correlations (37), or low communality (14). This yielded a final set of 13 items for inclusion in the OVHS.

EFA was performed on the UK sample. With the 13-item EFA, Bartlett's test of sphericity ($p<0.0001$) suggested that the data are appropriate for EFA; this was confirmed by the KMO measure of sampling adequacy, which was excellent at 0.90. Individual KMO values ranged from 0.70 to 0.96 for each item. A scree plot, minimum average partial correlation analysis, and parallel analysis all suggested a three-factor solution (Dinno, 2009; Garrido et al, 2011; Ledesma et al, 2015).

These three factors reflect domains of 1) vaccine beliefs (7 items), 2) vaccine-induced pain (3 items), and 3) a desire to engage in personal deliberation (3 items); domain names were based on the overall clinical theme of the items clustering in each domain. The eigenvalues for these factors were 6.10, 1.54, and 1.07 accounting for 70.07%, 17.71%, and 12.23% of the variance, respectively. Rotated factor loadings are included in Table 3.1. Factor intercorrelations are presented in Table 3.2, and the full OVHS, including scoring key, is outlined in Table 3.3.

	Beliefs	Pain	Deliberation
I would vaccinate a child in my care according to the standard vaccine schedule.	**0.9251**	-0.0188	-0.0462
The benefits of vaccines outweigh their risks.	**0.8972**	-0.0630	0.0505
Vaccines save lives.	**0.8806**	0.1093	-0.0830
It is unhealthy for children to be vaccinated as much as they are today.	**0.8720**	-0.0024	0.0705
I believe getting vaccinated is a social responsibility.	**0.8582**	-0.0565	0.0276
I would discourage my friends or family from getting vaccinated.	**0.7707**	0.1557	-0.0121
Drug companies prioritize profit over safety in vaccine development.	**0.6909**	-0.0742	0.1156
I have delayed getting a vaccine specifically because of the needle.	-0.1221	**0.9379**	0.0470
I have refused a vaccine in the past because of the associated pain.	0.2490	**0.7532**	-0.1195
Getting vaccinated is painful.	0.0237	**0.5617**	0.0873
It is important for me to personally understand the risks and benefits of vaccines before agreeing to be vaccinated.	-0.0045	-0.0877	**0.7909**
When considering vaccination, it is important to weigh the positives and negatives.	0.0005	-0.0152	**0.6941**
Doing my own research on vaccines is important to me.	0.0536	0.1303	**0.6775**

Table 3.1 Rotated factor loadings for the OVHS.

Factor	Beliefs	Pain	Deliberation
Beliefs	1		
Pain	.4327	1	
Deliberation	.4192	.1148	1

Table 3.2 Factor intercorrelations.

#	Item	Domain
1	I would vaccinate a child in my care according to the standard vaccine schedule.	Beliefs
2	The benefits of vaccines outweigh their risks.	Beliefs
3	Vaccines save lives.	Beliefs
4	It is unhealthy for children to be vaccinated as much as they are today.	Beliefs
5	I believe getting vaccinated is a social responsibility.	Beliefs
6	I would discourage my friends or family from getting vaccinated.	Beliefs
7	Drug companies prioritize profit over safety in vaccine development.	Beliefs
8	I have delayed getting a vaccine specifically because of the needle.	Pain
9	I have refused a vaccine in the past because of the associated pain.	Pain
10	Getting vaccinated is painful.	Pain
11	It is important for me to personally understand the risks and benefits of vaccines before agreeing to be vaccinated.	Deliberation
12	When considering vaccination, it is important to weigh the positives and negatives.	Deliberation
13	Doing my own research on vaccines is important to me.	Deliberation

Table 3.3 The OVHS.

All items are scored using a 5-point Likert scale ranging from disagree completely to agree completely. Higher scores reflect increasing hesitancy. Items 1, 2, 3, and 5 are reverse scored. The total score can range from 13 to 65.

3.3.3 Confirmatory Factor Analysis

To further assess the model developed in the EFA, a CFA was performed on the US sample. The confirmatory model created with structural equation modelling and a Satorra-Bentler scaled test statistic of the three-factor 13-item model demonstrated excellent goodness of fit characteristics, with RMSEA=0.044, SRMR=0.041, and CFI=0.977.

The median (IQR) OVHS score was 27 (24-33) in the UK sample and 28 (24-36) in the US sample (p=0.02). Descriptive statistics for each subscale are reflected in Table 3.4.

	UK Sample Median (IQR)	**US Sample Median (IQR)**	**p-value**
Beliefs	11 (8-15)	11 (8-17)	0.93
Pain	4 (3-6)	4 (3-6)	0.02
Deliberation	12 (10-14)	13 (11-14)	<0.001
OVHS Overall	27 (24-33)	28 (24-36)	0.02

Table 3.4 Comparison of scores between UK and US Samples.

The table presents the median and interquartile range (IQR) values for responses in the domains of Beliefs, Pain, Deliberation, and overall Oxford Vaccine Hesitancy Scale (OVHS) score. P-values are derived from the Mann-Whitney U test.

3.3.4 Scale Reliability and Validity

Reliability as internal consistency was measured using McDonald's omega; values for the overall scale were very good at 0.86 and 0.87 in the UK and US samples, respectively. Values for the three subscales are included in Table 3.5.

	UK Sample	**US Sample**
Beliefs	0.92	0.93
Pain	0.67	0.74
Deliberation	0.71	0.75
Overall OVHS	**0.86**	**0.87**

Table 3.5 Subscale and overall reliability (assessed using McDonald's Omega).

Predictive validity between the OVHS score and binary COVID-19 vaccination status highlighted the clinical value of the scale, with logistic regression odds ratios of

association (95% CI) of 0.07 (0.04, 0.13), p<0.0001 for the UK sample and 0.24 (0.17, 0.32), p<0.0001 for the US sample, for each standard deviation increase in OVHS score. That is, for each standard deviation increase in OVHS score, there was a 93% reduction in the odds of being vaccinated against COVID-19 in the UK sample and a 76% reduction in the odds of being vaccinated in the US sample.

Convergent validity between the OVHS score and the 5C short version scale demonstrated a correlation coefficient of 0.32 (p<0.0001), further bolstering our validity argument (Betsch et al, 2018).

Discriminant validity between an unrelated concept—enjoyment of outdoor activities—and OVHS score supported that the OVHS does not predict this unrelated outcome, with logistic regression odds ratios of association (95% CI) of 1.06 (0.88, 1.29), p=0.523 for the UK sample and 1.14 (0.93, 1.40), p=0.200 for the US sample for each standard deviation increase in OVHS score.

3.4 Discussion

Our work highlights the development and validation of the OVHS, a multidimensional instrument for assessing and quantifying vaccine hesitancy. Ultimately, despite the wide range of domains that were considered, we included only three broad top-level domains in the OVHS: 1) vaccine beliefs, encompassing concepts of efficacy, risk-benefit trade-off, planned behaviour, social responsibility, vaccine evangelism, and pharmaceutical company profiteering; 2) pain, reflecting the prominence of concerns regarding needle-induced pain in driving vaccine attitudes (Kantor et al, 2023); and 3) deliberation, or the concept that it is incumbent on each

individual to do their own research on vaccines, and that this personal deliberative process may contribute to vaccine acceptance or hesitancy.

3.4.1 Vaccine hesitancy

Vaccine hesitancy has shifted from an area of niche academic and scientific interest to not only a top-10 public health priority, but a socially trending public phenomenon as well (Mascherini & Nivakoski, 2022; Rathje et al, 2022). Indeed, while vaccine hesitancy is a multifactorial process, 'a motivational state of being conflicted about, or opposed to, getting vaccinated... includ[ing] intentions and willingness,'(World Health Organization, 2022) a recent systematic review of 422 studies addressing vaccine hesitancy highlighted the complexity of the concept, noting that three central approaches to understanding the phenomenon are reflected in the literature: cognition or affect, behaviour, and decision making. The authors suggested that the central theme of vaccine hesitancy is a 'state of indecisiveness' regarding the decision to vaccinate (Bussink-Voorend et al, 2022). This model—a state of being conflicted or indecisive regarding vaccination with the process driven by cognition and manifested as decision making—aligns broadly with the two OVHS domains of beliefs and deliberation.

3.4.2 Advantages of the OVHS

Several vaccine hesitancy scales have been created recently, but as highlighted previously (Betsch et al, 2018) most focus exclusively on a narrow range of domains, and those that take a more multidimensional approach often rely on a priori-

developed theoretical models—rather than qualitative research—in developing a range of domains for inclusion (Gilkey et al, 2014). Still other scales were designed for paediatric vaccines exclusively (Opel et al, 2011a; Opel et al, 2011b; Sarathchandra et al, 2018; Shapiro et al, 2018), and are therefore less appropriate for understanding responses to vaccination more generally.

Indeed, two of our three domains—pain and deliberation—represent a departure from prior scales and may provide additional actionable insight into the drivers of vaccine hesitancy. Understanding vaccine refusal is certainly helpful, but better appreciating if groups or individuals are on a path that may lead towards vaccine hesitancy at an earlier phase has critical public health implications and has not been previously explored. By including a domain addressing pain, our goal was to ensure that the OVHS included not only immutable predictors or antecedents of behaviour, but potentially modifiable associations as well. This is particularly important given that scales such as the OVHS may be used not simply to assess public opinion on a longitudinal or cross-sectional basis, but also to aid in the development of alternative approaches and technologies for vaccination promotion or delivery (Hettinga & Carlisle, 2020; Kantor et al, 2023). By addressing modifiable risk factors for vaccine hesitancy, such as pain, the OVHS may be particularly well-suited to future public health and interventional studies evaluating ways in which vaccine hesitancy can be not only studied and quantified, but ameliorated as well.

Similarly, the deliberation domain addresses concerns regarding contestable science and social context that have proliferated in the past decade, and particularly since the COVID-19 pandemic, making the concept of vaccine hesitancy far more

nuanced and heterogeneous than previously appreciated (Attwell et al, 2018; Carrion, 2017; Duchsherer et al, 2020; Wang et al, 2015; Ward et al, 2017). Indeed, deliberation has been highlighted as a precursor to the distrust and dissent that are the hallmarks of vaccine refusal (Tram et al, 2022), and our scale is the first to include this domain that could potentially serve as an early indicator of future vaccine refusal.

3.4.3 Study strengths

Our study has several important strengths. First, we used a combination of an evidence-based literature review coupled with qualitative research to develop the domains and items of interest. This approach helps maximize our ability to capture domains and items of interest, particularly given the historical reliance on a priori defined theoretical frameworks with prior scales. Second, our early qualitative research relied on both the general population and vaccine hesitant respondents, a unique approach that provided useful insights into priorities and concerns for the potentially marginalized vaccine hesitant population. Third, our work relied on large demographically representative samples from both the US and UK, further bolstering our generalizability and the potential clinical utility of the scale. Fourth, we relied on robust statistical approaches developed for Likert-scale ordinal data, such as EFA using a polychoric factor matrix, to mitigate concerns regarding the appropriate statistical approach to Likert-scale data (DeWees et al, 2020; Wu & Leung, 2017). Fifth, the outstanding predictive validity of the OVHS—a single standard deviation increase in score is associated with a 76% to 93% reduction in the odds of having been vaccinated against COVID-19—strongly bolsters our validity argument by highlighting the real-

world association between OVHS score and vaccine refusal, further supported by our convergent validity with the 5C short form scale. Sixth, the baseline factor intercorrelations suggest that the subscales measure similar, but not identical, concepts, further supporting our validity argument, though the lower correlation between deliberation and pain (r=0.11) suggest that these particular two domains may capture independent contributors to vaccine hesitancy. Seventh, baseline descriptive statistics, such as median OVHS subscale scores, were similar in the UK and US populations, though small (and clinically insignificant) absolute differences did reach statistical significance, likely due to the large sample sizes used in our validation cohorts. Nevertheless, future studies evaluating the OVHS in large and varied populations will be helpful. Finally, the OVHS was designed for use in both adult and paediatric populations, and questions were worded to apply to both adults contemplating their own vaccine choices and parents deciding on the best approach for their children to maximize the utility of our scale for future research studies.

3.4.4 Study limitations

This study also has important limitations. First, survey-panel participants may differ in important ways from the general population, though the platform used in our study has been used widely, our large demographically representative samples from both the US and UK likely mitigate these issues, and the robustness of our findings across populations suggests that our generalizability is indeed adequate. Second, we included unequal numbers of items across domains, with the beliefs domain accounting for seven of the 13 items in the scale. Given the breadth of this concept,

however, which accounts potentially for several second-order domains, and the meaningful conceptual and statistical partitioning of items across domains, this is unlikely to be a major limitation. Third, the reliability coefficients for the pain and deliberation subscales ranged from 0.67 to 0.75; that said, McDonald's Omega is sensitive to the number of items in each subscale, and indeed for shorter subscales reliability coefficients likely underestimate the true reliability (Eisinga et al, 2013). Fourth, qualitative data were collected through a series of open- and closed-ended survey questions provided to both a convenience sample and a sample composed of respondents that had never been vaccinated against COVID-19; it is possible that in-person interactive focus groups could have provided additional insights into these responses, though an in-person approach may have constrained the number of vaccine hesitant respondents that could have been feasibly included in the qualitative portion of the study as such subjects are often challenging to recruit (Hilário et al, 2023). Finally, while this manuscript presents the development and initial validity evidence and reliability testing for the OVHS, future work may be helpful to better characterize the scale across differing cultural populations, perform test-retest reliability assessments, and conduct invariance testing.

3.4.5 Conclusions

While vaccine hesitancy is a critical priority for public health investigation, it is also challenging to study reproducibly. We therefore developed the OVHS to address the real-world observed phenomenon of vaccine hesitancy, particularly as it contributes to vaccine refusal, by approaching the challenge pragmatically and relying

on not only a theoretical model of hesitancy but qualitative data as well. By yielding a simple three-factor model reflecting vaccine beliefs, concerns regarding pain, and the need for a deliberative process, the OVHS hopefully represents a clinically useful instrument with the potential to aid in both cross-sectional and longitudinal assessments of hesitancy. The OVHS may also provide insight into how the myriad contributors to hesitancy interact and affect behaviour while ultimately assisting in developing alternative technological approaches to vaccination that may broaden access to vaccines worldwide.

Chapter 4: Discriminative cut-offs, concurrent criterion validity, and test-retest reliability of the Oxford Vaccine Hesitancy Scale

4.1 Introduction

Validated scales help investigators and clinicians better understand contributors to vaccine hesitancy, and such scales can be used to help identify those likely to refuse vaccination, permitting tailored outreach and intervention (Kantor et al, 2024b; Martin & Petrie, 2017a; Opel et al, 2013; Shapiro et al, 2018). Yet while hesitancy exists on a continuum—and indeed that is part of the value-add of using such scales—it is also important to quantify how effectively a given scale discriminates between those who do and do not exhibit vaccine hesitant behaviours of interest, what discrete scale cut-off points are associated with historical or future behaviours, and how reliably a scale performs when provided to the same respondent at a later timepoint (Betsch et al, 2018; Ghazy et al, 2021; Machida et al, 2023).

Scale discrimination has classically been demonstrated through the receiver operating characteristic (ROC) curve, and particularly by presenting the area under the curve (AUC) as a measure of discriminatory ability (Hanley & McNeil, 1982; Zweig & Campbell, 1993). Yet clinicians and researchers may benefit from a single operating cut-off point that is transparently derived and is transportable across populations and time. Since most vaccine hesitancy scales present data in a continuous fashion, this leaves cut-off choice (if any) ad hoc and vulnerable to prevalence shifts.

Some vaccine hesitancy scales do have established thresholds. For example, the Parent Attitudes about Childhood Vaccines (PACV) scale uses a reported cut-off of 50, which predicts under-immunisation across settings (Opel et al, 2013). Other

instruments also have empirically derived cut-offs, including the 5C (Ghazy et al, 2021), the 7C (Machida et al, 2023), and a single-item VHA-10 (Bussink-Voorend et al, 2025). Still, these cut-offs were often developed on niche populations and, importantly, their clinical utility is limited by any shortcomings of the underlying scales (Kantor et al, 2024b).

Similarly, most scales are validated at a single timepoint in a cross-sectional fashion, but understanding how stable scale scores are over time—assuming there has been no change in underlying attitudes—helps support scale reliability and has important real-world implications in terms of clinical utility (Terwee et al, 2007). While select vaccine hesitancy scales—such as the Malay and Turkish versions of the PACV (Abd Halim et al, 2020; Mutlu et al, 2023), the French and Urdu versions of the Vaccination Attitudes Examination (VAX) (Eisenblaetter et al, 2023; Wilhelm et al, 2024), the Polish version of the Vaccine Hesitancy Scale (VHS) (Duplaga et al, 2022), the Turkish version of the Oxford COVID-19 Vaccine Hesitancy Scale (Karabulut et al, 2022), and the Korean version of the Psychological Antecedents of Vaccination (5C) scale (Bae & Kim, 2023)—have been assessed for test–retest reliability, reported intraclass correlation coefficients (ICCs) vary widely from 0.53 to 0.95 depending on the instrument, adaptation, and domain. Many of these temporal stability studies have been conducted on small, specialized populations, and indeed some have noted that test-retest reliability is infrequently performed in vaccine hesitancy scale assessments (Shapiro et al, 2021).

The Oxford Vaccine Hesitancy Scale (OVHS) is a 13-item validated multidimensional instrument where scores range from 13 to 65, with higher scores

suggesting higher levels of vaccine hesitancy (Kantor et al, 2024b). While initial scale development and validation work suggested its strong psychometric properties, multidimensional nature, and cross-population validity and robustness, it was initially presented as a continuous scale, while test-retest reliability—the temporal stability of scores—was not assessed.

We therefore developed and implemented a scale-agnostic methodology to a) assess the discrimination of the OVHS when evaluating self-declared historical COVID-19 vaccine refusal as an outcome; b) derive and test the performance of a single operational cut-off point for this outcome; and c) assess the test-retest reliability of the scale. While we used COVID-19 vaccine refusal as an example, future studies may explore whether our cut-off points remain stable when evaluating the discrimination of the OVHS in predicting vaccine refusal in other contexts and populations.

4.2 Methods

4.2.1 Samples

This study included five datasets (including one repeated assessment) on four separate demographically representative samples in the UK and US; details of our sampling strategy have been previously reported (Kantor & Kantor, 2021; Kantor et al, 2024a; Kantor et al, 2024b; Kantor et al, 2025; Kantor et al, 2023). Surveys were developed using Qualtrics (Qualtrics International, Provo, Utah, USA) and we used a survey panel approach via the academic survey site Prolific Academic (Oxford, UK).

For the UK datasets, demographically representative samples were stratified by age, sex, and ethnicity (based on data from the UK Office of National Statistics). We assessed data from two distinct samples, $UK_{1.1}$ and UK_2; the $UK_{1.1}$ sample was re-

assessed six weeks after initial survey administration to assess test-retest reliability. Our two US samples, US_1 and US_2, were stratified by age, sex, and ethnicity based on 2015 US national census bureau data. Since prevalence and timing differed by dataset, we did not pool data.

This study was approved by the University of Oxford Medical Sciences Interdivisional Research Ethics Committee (approval reference R81585/ RE001). Study participants were reimbursed less than £5/$5 for their time. Participants provided informed consent and were permitted to withdraw from the study at any time. Demographic details were self-declared and linked to responses using a unique 25-character alphanumeric identifier which was also used for clustering when assessing test-retest reliability. Multiple attention check items were included in the survey instrument, and responses were excluded from the analysis if a respondent failed two or more attention checks. Analyses included only respondents with complete OVHS scores and outcomes. Psychometric term definitions are included in the Appendix.

Statistical analyses were performed using Stata for Mac version 16 (Stata Corporation, College Station, Texas, USA).

4.2.2 Measures

The OVHS includes a total of 13 items across three domains, yielding a possible score range of 13-65, with higher scores reflecting greater vaccine hesitancy (Kantor et al, 2024b). The theoretical neutral midpoint of this scale is thus 39, reflecting a mean item score of 3.

The reference standard for assessing operating characteristics was defined a priori as a self-declared history of receiving no vaccinations against COVID-19, which was treated as a binary variable for unvaccinated status. Dataset-specific rates of history of COVID-19 vaccine refusal were used to calculate positive and negative predictive values (PPV and NPV).

4.2.3 ROC curves

We generated an ROC curve for each sample and calculated the AUC. We assessed 95% bootstrap confidence intervals for the AUC using 10,000 resamples using both bias-corrected and accelerated (BCa) and percentile methods. Resampling was stratified by outcome (refusal vs non-refusal) and in the $UK_{1.2}$ sample was clustered by participant ID to ensure within-person correlation.

4.2.4 Cut-off derivation algorithms

We developed an algorithm to assess an ideal discriminatory cut-off taking advantage of our multiple samples. Let datasets be indexed by $\mathrm{k} = 1, \ldots, \mathcal{K}$, integer cut-offs by $\mathcal{C}$, sensitivity/specificity at cut-off c by $Se_{k(c)}$, $Sp_{k(c)}$, and Youden's index J by $\mathrm{J_k(c)} = \mathrm{Se_k(c)} + \mathrm{Sp_k(c)} - 1$. Using the $UK_{1.1}$ dataset, *d*, we examined all possible integer thresholds to identify a near-optimal Youden's J plateau around the global maximum, $J_d^{max} = max_{c \in C} J_d(c)$ with $P = \{\, c \in \mathcal{C} : J_d(c) \geq J_d^{max} - \varepsilon \,\}$ and $\varepsilon = 0.01$ (as we defined near-optimal as ΔJ≤ 0.01 from the highest J). From this plateau we chose the smallest threshold that satisfied a cross-sample specificity floor of ≥ 0.75 across the majority of samples—a pragmatic approach designed to prioritize

sensitivity while limiting false positives, to form the admissible set $\mathcal{A} = \{ c \in \mathcal{P}: \frac{1}{K}\sum_{k=1}^{K} 1\{ Sp_k(c) \geq \tau \} > \alpha\}$ using $\tau = 0.75$ as a specificity floor and $\alpha = \frac{\lceil K/2 \rceil}{K}$ so that the floor would be met in the majority of datasets. We then selected the transportable cut-off as the smallest admissible threshold $c^\star = \min \mathcal{A}$, which yielded OVHS≥35. After fixing $c = c^\star$, we evaluated transportability in the remaining datasets by reporting operating characteristics at $c^\star$ and, for context, each dataset's J-maximizing threshold; where thresholds tied on $J_k(c)$ within a dataset, we report the lower c, favouring sensitivity.

A secondary ROC-based criterion was assessed using the Index of Union (IU) rule (Ünal et al, 2025); we then expanded on the published approach applicable to only a single dataset and created a novel algorithm developing an aggregate IU (IU_{agg}). Briefly, for dataset k with AUC $A_k$and sensitivity/specificity at cut-off c given by $Se_{k(c)}$, $Sp_{k(c)}$, we computed $IU_{k(c)} = \left|Se_{k(c)} - A_k\right| + \left|Sp_{k(c)} - A_k\right|$ selecting c that minimized $IU_{k(c)}$ with ties broken by minimizing $\left|Se_{k(c)} - Sp_{k(c)}\right|$. Since we sought a single transportable cut-off, we then defined a novel aggregate IU target, $IU_{agg(c)} = \sum_k w_k IU_{k(c)}$ with equal weights, w_k. The universal IU cut-off was the c minimizing $IU_{agg(c)}$ across $c \in \{34, \ldots, 38\}$.

To further visualize threshold behaviour, we generated a multiple sample two-graph ROC (TG-ROC) style plot by plotting sensitivity and specificity against the OVHS integer cut-off for each dataset (Greiner, 1996). We included all available integer cut-offs from 32 to 40, connected observed points without smoothing or interpolation, colour-coded datasets, used dashed lines for sensitivity and solid lines for specificity, and marked the fixed operating cut-off with a vertical reference at 35.

4.2.5 Test-retest reliability

Test-retest reliability was performed using intraclass correlation coefficients (ICCs) with a two-way, mixed-effects, absolute-agreement single-measure model. This is denoted ICC(A,1) in the Shrout–Fleiss framework (Shrout & Fleiss, 1979), where 'A' denotes absolute agreement (penalizing both random error and systematic shift between administrations) and '1' denotes a single-measure estimate. We report ICCs for the OVHS overall and each subscale in the UK population measured across two timepoints six weeks apart. The test–retest analysis included only participants with complete OVHS at both timepoints (n = 462). As a secondary analysis, we also report the average-measure ICC(A,2), which reflect the reliability of the mean of the two OVHS administrations.

4.3 Results

4.3.1 Sample Population

Our analyses incorporated data from 2,451 assessments on 1,989 demographically representative respondents in the UK and US. Mean ages across samples ranged from 45 to 47, and the age range of respondents was from 18 to 85. The demographic mix—age, sex, and race—of the samples reflected UK or US population-level distributions; details on these samples have been described previously (Kantor et al, 2024b; Kantor et al, 2025). Because samples were independent, we did not pool data for analysis. Observed COVID-19 vaccine refusal prevalence ranged from 8.4-20.2%.

4.3.2 ROC Curves and AUC values

The ROC curves for the $UK_{1.1}$ (derivation) and $UK_{1.2}$ (retest) datasets are shown in Figures 4.1 and 4.2. The AUC for the derivation sample was 0.960 (95% CI 0.890–0.981); across all samples, the AUC was highest in the UK retest sample ($UK_{1.2}$) and lowest in US_2 (Table 4.1). AUCs with 95% CI for the OVHS across all samples are included in Table 4.1 and range from 0.760 to 0.971. For reference, in our derivation sample the AUC for the 5C short form scale (Betsch et al, 2018) was 0.752.

Sample	**n**	**AUC**	**95% CI (BCa)**	**95% CI (Percentile)**
$UK_{1.1}$	504	0.960	0.890–0.981	0.920–0.986
$UK_{1.2}$	462	0.971	0.951–0.983	0.954–0.984
UK_2	491	0.891	0.844–0.926	0.848–0.929
US_1	500	0.851	0.799–0.891	0.803–0.894
US_2	500	0.760	0.704–0.807	0.706–0.809

Table 4.1 Discriminative performance (ROC AUC) of the OVHS for classifying historical COVID-19 vaccination refusal across samples.

Note: AUCs are shown with 95% bias-corrected and accelerated (BCa) bootstrap CIs (primary) and 95% percentile bootstrap CIs; all CIs are calculated using 10,000 bootstrap replications with resampling stratified by outcome (unvaccinated status) and clustered by participant ID. $UK_{1.2}$ denotes the second UK test–retest administration (~six weeks after $UK_{1.1}$).

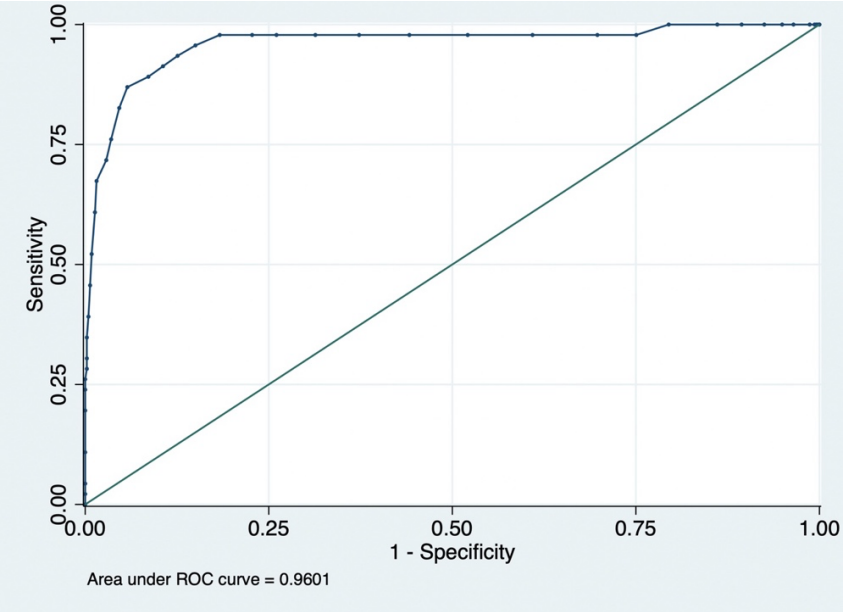


Figure 4.1 ROC curve for the OVHS in the $UK_{1.1}$ dataset.

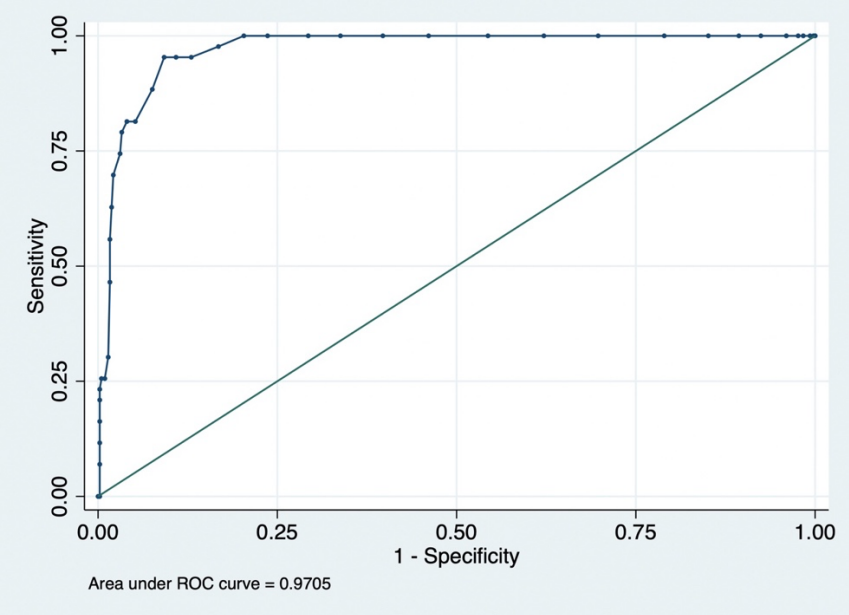


Figure 4.2 ROC curve for the OVHS in the $UK_{1.2}$ dataset.

4.3.3 Cut-Off Derivation

In the $UK_{1.1}$ derivation set, Youden's J exhibited a near-optimal plateau spanning OVHS scores of 34–38, where the ΔJ was ≤ 0.01, with the numerical maximum at ≥38 (J = 0.812). Based on our multiple dataset selection algorithm, we

chose ≥35 as the fixed operating cut-off as it lies on this plateau and met the predefined specificity floor criterion of ≥0.75 across the majority of datasets. A cross-dataset Youden heatmap highlighting this plateau is included in Figure 4.3.

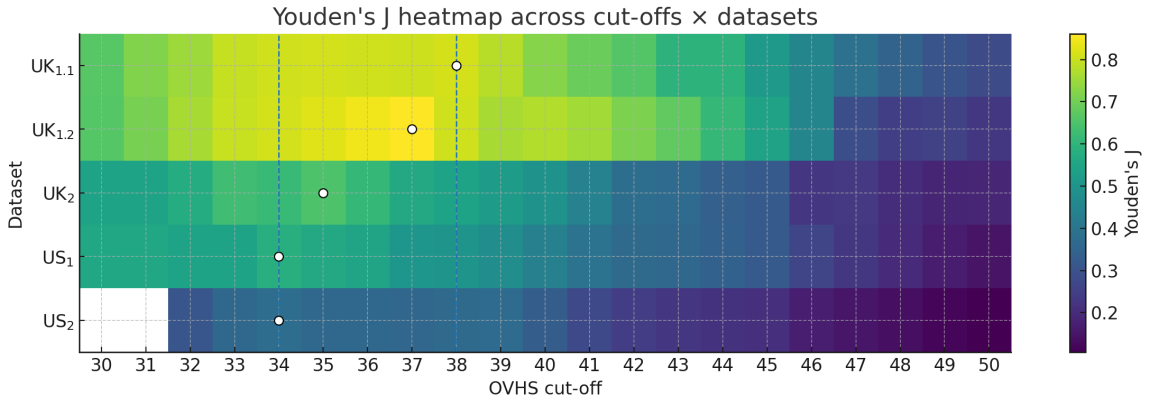


Figure 4.3 Heatmap demonstrating the near-optimal plateau.

Dashed lines indicate the plateau boundaries while circles note dataset-specific J maxima.

For our secondary analysis using the IU rule to derive a cut-off, the dataset IU minima were ≥38 ($UK_{1.1}$), ≥37 ($UK_{1.2}$), ≥36 (UK_2), ≥34 (US_1), and ≥36 (US_2). IU_{agg} was minimized at OVHS ≥35 so that $IU_{agg}(35) = 0.6626$; the next best cut-off was 36 with $IU_{agg}(36) = 0.6986$ (Table 4.2). This optimal ≥35 cut-off was consistent with our Youden plateau–based choice and prespecified transportability criteria.

Cut-off	**34**	**35**	**36**	**37**	**38**	**IU minimum**
$UK_{1.1}$	0.1136	0.1110	0.1130	0.1148	**0.1078**	38
$UK_{1.2}$	0.1036	0.1185	0.0972	**0.0807**	0.1340	37
UK_2	0.2657	0.2319	**0.1730**	0.2203	0.2378	36
US_1	**0.1195**	0.1401	0.1550	0.2210	0.2607	34
US_2	0.2469	0.1598	**0.1414**	0.1639	0.1536	36
Aggregate	0.7902	**0.6626**	0.6986	0.8007	0.8939	35

Table 4.2 Index of Union (IU) across candidate OVHS cut-offs (34–38) by dataset, with per-dataset IU minimum and aggregate IU included.

Minima are bolded for reference; these values correspond to cut-offs which provide the best discrimination for these samples.

Using a single cut-off of ≥35, classification performance was strong across all UK datasets and acceptable across the US datasets; the NPV of the ≥35 cut-off was >0.9 across all samples (Table 4.3). For context, Table 4.4 reports operating characteristics of each dataset's maximum J with ΔJ relative to the ≥35 cut-off.

Metric	**$UK_{1.1}$**	**$UK_{1.2}$**	**UK_2**	**US_1**	**US_2**
Prevalence	0.0922	0.0923	0.0840	0.1805	0.2016
Sensitivity	0.9348	0.9535	0.8780	0.7416	0.7327
Specificity	0.8742	0.8700	0.7720	0.8193	0.6275
Youden J	0.8090	0.8235	0.6500	0.5609	0.3602
Balanced accuracy	0.9050	0.9118	0.8250	0.7800	0.6801
LR +	7.4000	7.3332	3.8480	4.1040	1.9669
LR -	0.0750	0.0535	0.1580	0.3154	0.4260
PPV	0.4301	0.4272	0.2610	0.4748	0.3319
NPV	0.9925	0.9946	0.9857	0.9350	0.9029

Table 4.3 Discriminative performance of total OVHS score≥35 for classifying self-reported historical COVID-19 vaccine refusal across five samples.

Prevalence reflects the prevalence of past COVID-19 vaccine refusal (a history of never receiving a single vaccine) by self-report. Note: $UK_{1.2}$ includes a subset of the same subjects as in $UK_{1.1}$ at a timepoint six weeks later.

Metric	**$UK_{1.1}$**	**$UK_{1.2}$**	**UK_2**	**US_1**	**US_2**
Optimal OVHS Cut-off (≥)	38	37	35	34	34
Prevalence	0.0922	0.0923	0.0840	0.1805	0.2016
Sensitivity	0.8696	0.9535	0.8780	0.7978	0.8119
Specificity	0.9426	0.9078	0.7720	0.7847	0.5650
Youden J	0.8122	0.8613	0.6500	0.5825	0.3769
ΔJ to ≥ 35	0.0032	0.0378	0.0000	0.0216	0.0167

Balanced accuracy	0.9061	0.9306	0.8250	0.7913	0.6885
LR +	15.1498	10.3416	3.8509	3.7055	1.8664
LR -	0.1383	0.0512	0.1580	0.2577	0.3329
PPV	0.6061	0.5126	0.2610	0.4494	0.3203
NPV	0.9861	0.9948	0.9857	0.9463	0.9225

Table 4.4 Operating characteristics of sample-specific optimal OVHS cut-offs based on maximizing Youden's J for each sample.

As seen in Figure 4.4, sensitivity decreases and specificity increases with higher cut-offs in each cohort, with cross-over points aligning with tabulated operating characteristics (Table 4.4). Figure 4.3 also highlights the near-optimal region, showing that the fixed operating cut-off ≥35 lies within the stable range while transporting reasonably across all five datasets.

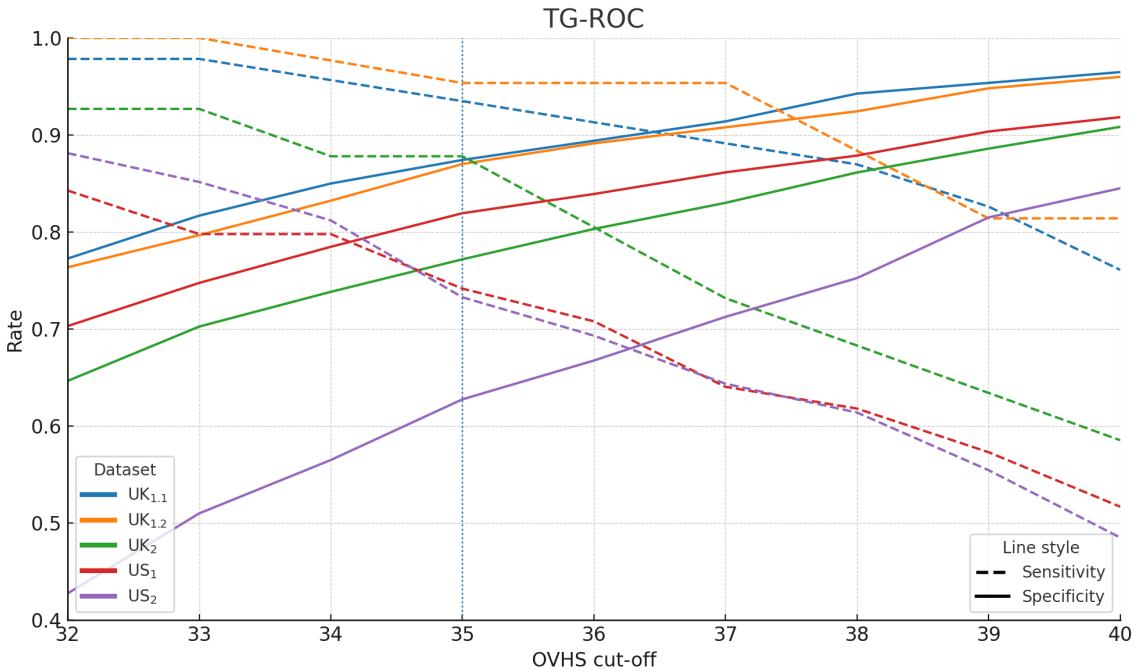


Figure 4.4 Two-graph ROC-style figure.

Sensitivity (dashed) and specificity (solid) vs OVHS cut-off across five datasets; vertical line denotes the fixed operating cut-off (≥35).

4.3.4 Test-retest reliability

With a two-way mixed-effect absolute-agreement single-measure model, the overall OVHS demonstrated ICC(A,1)=0.884 (95% CI 0.863, 0.903) in the n=462 two-timepoint sample measured six weeks apart. Subscales showed ICC(A,1) values from 0.649 to 0.901, with lower values seen for shorter three-item subscales; average-measure reliability was ICC(A,2)=0.939, with subscales ranging from 0.787 to 0.948 (Table 4.5).

Scale		**ICC**	**95% CI**
OVHS overall (13 items)	Single measure	0.884	0.863–0.903
	Average measure	0.939	0.926–0.949
Beliefs (7 items)	Single measure	0.901	0.882–0.916
	Average measure	0.948	0.937–0.956
Deliberation (3 items)	Single measure	0.649	0.593–0.699
	Average measure	0.787	0.745–0.823
Pain (3 items)	Single measure	0.677	0.624–0.724
	Average measure	0.807	0.768–0.840

Table 4.5 Two-way mixed-effects absolute-agreement single- and average-measure ICC(A,1) and ICC(A,2) model results, with n=462 over two timepoints six weeks apart in a UK sample.

All p-values <0.001.

4.4 Discussion

Across 2,451 assessments on 1,989 demographically representative respondents in the UK and US, we found that a) the OVHS demonstrated good to excellent discrimination (ROC AUC in the derivation sample was 0.960); b) a cross-population cut-off of ≥35 is ideal, yielding good to excellent operating characteristics across datasets; and c) the OVHS demonstrates very good to excellent temporal

stability in a UK sample with (ICC(A,1)=0.884; ICC(A,2)=0.939). The near-midpoint clustering of optimal operating points (34–38, with a theoretical neutral midpoint of 39) also supports the construct validity of the OVHS, as the probability of COVID-19 vaccine refusal rises as item responses shift toward increased hesitancy, even below the neutral mean-item score. While the AUC was lower in the US samples, this reflects expected changes based on case mix and supports that the model was not overfitted. Finally, while we considered using separate cut-offs for the UK and US, the small ΔJ gains that could be seen with using two cut-offs are offset by the benefit and simplicity of a single global cut-off in these populations. Together, these data support the validity and reliability of the OVHS while also providing actionable data that could be helpful in both clinical and research contexts.

We introduce several methodological nuances that may be useful for other research groups in reporting discrimination, highlighting concurrent criterion validity, and evaluating ideal cut-off point across datasets more broadly. We include a matrix table of AUCs using computationally intensive bootstrapped confidence intervals over 10,000 resamples using both BCa and percentile methods to provide a conservative and realistic assessment of AUC values. We then used a novel Youden's J-based algorithm to locate a near-optimal plateau of integer scores and select a cut-off by applying a small-change rule ($\Delta J \leq 0.01$) within that plateau on the derivation dataset; from the admissible set we chose the smallest threshold that satisfied a cross-dataset specificity floor of ≥ 0.75, thereby aggregating results while also prioritizing sensitivity and high NPV for screening while limiting false positives and improving transportability. To make selection logic transparent across cohorts, we also present a

single threshold by dataset heatmap, a compact, reproducible, and novel visual for cross-cohort comparisons. Finally, we performed a secondary, ROC-native assessment using the IU and we introduce an algorithm to derive the IU_{agg} as a prevalence-agnostic aid to cut-off derivation across multiple samples, confirming that the selected threshold remains near the global optimum while sensitivity and specificity are targeted toward each dataset's AUC.

We adopted a conservative approach to test–retest reliability by estimating a two-way mixed-effects, absolute-agreement, single-measure intraclass correlation—ICC(A,1) in the Shrout–Fleiss framework (Liljequist et al, 2019; Shrout & Fleiss, 1979)—over the six-week interval. This model treats participants as random and the two administrations as fixed, penalizing any systematic shift in means between timepoints—not only random error. This makes this approach the preferred lower-bound estimate when the goal is to judge whether a single administration at a second timepoint can be used interchangeably with the first. Confidence intervals were obtained from the F distribution, and interpretation followed standard guidance (<0.50 poor, 0.50–0.75 moderate, 0.75–0.90 good, >0.90 excellent) (Koo & Li, 2016). Because programmes may sometimes average repeated administrations (e.g., for baseline screening or research composites), we also report the average-measure ICC—ICC(A,2)—which quantifies the reliability of the mean of the two administrations; as expected from the Spearman–Brown relation (de Vet et al, 2017), ICC(A,2) exceeds ICC(A,1), reflecting the precision gained by averaging. The six-week interval was chosen to reduce recall carry-over while remaining short enough to minimize true attitude change.

4.4.1 Study strengths

This study has several other strengths. First, while many studies evaluating cut-offs and test-retest reliability focus on small samples as part of a cross-cultural validation enterprise, we include data on almost 2,000 distinct respondents using demographically representative samples in the US and UK. We also evaluated performance separately across multiple independent datasets spanning countries and timepoints rather than pooling results, mirroring how programmes face prevalence variation (as highlighted in Figure 4.4, the US_2 dataset reflects such variation, and yet our cut-offs remain consistent). Second, our ≥35 cut-off was selected from a clearly documented flat optimum and fixed before transport testing, reducing the risk of overfitting. Third, discrimination was quantified with bootstrapped CIs, operating characteristics were presented in decision-relevant terms, and the heatmap and IU_{agg} provide reproducible, cross-cohort views of threshold behaviour. Fourth, temporal stability (test-retest reliability) was addressed directly using absolute-agreement ICCs, documenting excellent total-score reliability and moderate-to-excellent subscale stability as well—an aspect often omitted in reliability testing.

4.4.2 Study limitations

Our work also has some important limitations. First, we used self-reported historical COVID-19 vaccine refusal as the outcome of interest on which we assessed scale discrimination performance. Therefore, our data are subject to reporting biases and may not capture the appropriate threshold when generalized to other vaccination campaigns. Any threshold work is dependent outcome definition, however, though

future studies evaluating other outcomes will be beneficial. Second, PPV/NPV are prevalence-dependent; while NPV remained uniformly high at our cut-off of ≥35, PPV may vary as the cut-off is applied in settings with different rates of baseline hesitancy. Third, the test-retest analysis was conducted in the UK only, and over a six-week interval; longer time horizons and additional cross-national studies may help better characterize temporal stability. Fourth, while we provide ample evidence for the value of novel approaches to assessing validity such as our ΔJ algorithm, heatmap presentations, and the IU_{agg}, these are not established approaches. We therefore present these data as secondary exploratory analyses and as confirmation of more established ROC curve, AUC, and Youden's J-based cut-off assessments. The concordance between these approaches highlights the value of our approach and suggests it may be useful for other investigators in the future. Fifth, although our survey panel sampling was broadly demographically representative and included attention checks, the population of participants who register for survey panels may differ systematically from the general population or the population of individuals seeking medical care. This limitation, however, is shared by most sampling approaches, and the large size and demographic representativeness of our samples, with generally consistent cross-dataset results, further supports the robustness of our findings. Still, additional studies in different populations may be of value. Sixth, the PPV ranged from 0.26 to 0.47 and is inherently constrained by the relatively low prevalence of vaccine refusal in our samples (8.4–20.2%). In screening contexts with low base rates, even instruments with high sensitivity and specificity will yield modest PPV; these values are broadly comparable to those reported for other vaccine

hesitancy scales at similar prevalence levels. The uniformly high NPV (>0.90) suggests that the OVHS at ≥35 is effective for ruling out vaccine refusal. Finally, the OVHS uses a classical summed-score approach so that domains with more items contribute proportionally more to the total score. While this is intentional—the beliefs domain captures the broadest range of hesitancy determinants—it also means that conceptually overlapping items within a domain may up-weight that dimension relative to shorter subscales. Item response theory (IRT) models offer an alternative scoring framework in which each item is weighted by its discriminative ability and difficulty rather than by simple count, potentially yielding a more measurement-efficient composite. Future work exploring IRT-based scoring of the OVHS may be worthwhile.

4.4.3 Conclusions

This study adds to the validity and reliability evidence supporting the use of the OVHS as a feasible, multidimensional scale for assessing vaccine hesitancy. Although COVID-19 refusal served as the worked example, our derivation framework is scale- and endpoint-agnostic so that the same steps—Youden plateau identification subject to a cross-dataset specificity floor, heatmap visualization, and confirmation using IU_{agg}—can be applied to derive vaccine-agnostic or vaccine-specific cut-offs whenever a dichotomous reference is available. Indeed, these approaches could be explored more broadly when studying any algorithmic approach to cut-off derivation and have implications not only for epidemiologic research but for machine learning and artificial intelligence models where algorithmic approaches draw iteratively on

diverse datasets to develop optimal decision cut-offs. Future studies addressing different outcomes of interest using cross-cultural and -language approaches may be helpful to broaden the validity evidence for the OVHS even further.

Chapter 5: Development and Validation of the Oxford Benchmark Scale for Rating Vaccine Technologies (OBSRVT)

5.1 Introduction

Developing novel vaccine delivery technologies is an international public health priority, driven by both the ongoing challenge of vaccine hesitancy and the need for more efficient, deployable, and innovative approaches to vaccine delivery (Kanekiyo et al, 2019; Wallis et al, 2019). While traditional needle and syringe-based vaccination has represented the standard of care for decades, next-generation vaccine delivery technologies such as nasal sprays, microneedle patches, oral vaccines, and ultrasound are increasingly being investigated as alternative, and potentially superior, approaches to vaccine delivery (Hettinga & Carlisle, 2020; Hettinga et al, 2023; Kumar et al, 2022a; Rodgers et al, 2019; Tang et al, 2022).

Understanding public willingness to consider the uptake of next-generation vaccine delivery technologies is important for two central reasons. First, alternative vaccine delivery technologies may help overcome some of the technical and logistical barriers to vaccination—cold storage, enhancing immune response, risk of sharps injuries, and cost—that could lead to more seamless deployment (Hettinga & Carlisle, 2020; Hettinga et al, 2023). Second, these approaches may help mitigate some of the hesitancy surrounding vaccination, a phenomenon that may present an even greater, or at least more intransigent, barrier to effective vaccine delivery than deployment and efficacy challenges (Freeman et al, 2022).

Yet despite decades of work in developing alternatives to needle and syringe-based vaccination, public attitudes to these individual approaches remains unknown,

and little has been done to develop a validated approach to assessing patient and public preferences for novel vaccine delivery technologies (Kaaijk et al, 2014). Furthermore, novel technologies deployed in a medical setting are also sometimes themselves associated with increased mistrust (Caplan et al, 2024; Jaiswal & Halkitis, 2019; Nguyen et al, 2021a). Therefore, while developing needle-free approaches to vaccine delivery may represent a powerful modulatory tool in the armamentarium of public health researchers to mitigate vaccine hesitancy, it remains difficult to determine the net value of novel delivery technologies from the public perspective.

As new vaccine delivery approaches are considered, developed, and trialled, it will be crucial to provide researchers with validated instruments so that the preferences of the public and trial participants can be studied, considered, and addressed. On a practical level, while funding agencies and pharmaceutical companies are highly invested in developing new delivery approaches, a validated scale to assess acceptance and attitudes to specific approaches of interest could help with de-risking these technologies at an early stage, helping to target investment and research across the timeline of technological development (Abou-El-Enein et al, 2017).

We therefore sought to develop and validate a feasible, psychometrically sound, and multidimensional benchmark scale—the Oxford Benchmark Scale for Rating Vaccine Technologies (OBSRVT)—that could be used at all phases of alternative vaccine delivery approach development to gauge the relative preferences of the public for these technologies. Ultimately, such a scale may provide evidence of the putative benefits of novel, and sometimes costly, delivery technologies from a societal standpoint that could help tailor future development to better align with public

preferences while also helping to justify the research and development of these technologies. This scale will hopefully lead to a better understanding of the barriers to vaccine programme delivery in all its myriad forms, ultimately resulting in enhanced vaccine coverage and improved public health outcomes.

5.2 Methods

5.2.1 Data Collection, Population, and Overall Approach

Our overall methodology was similar to our prior work in developing instruments to assess needle attitudes and vaccine hesitancy as outlined in the preceding chapters (Kantor & Kantor, 2021; Kantor et al, 2024b; Kantor et al, 2023). Initially, we developed a series of questionnaire-based surveys in Qualtrics (Qualtrics International, Provo, Utah, USA) with responses solicited using a survey panel design through Prolific Academic (Oxford, UK) from two separate 50-member cohorts. The first cohort was a convenience sample and the second represented a vaccine hesitant group who self-identified as never having received a vaccination for COVID-19.

After this initial exploratory study, two separate 500-respondent demographically representative samples from the UK and US (based on UK Office of National Statistics or 2015 US national census bureau data, respectively, and matched by stratification on age, sex, and ethnicity) were targeted to develop the OBSRVT scale. The UK data were used for the exploratory factor analysis (EFA), while US data were used for confirmatory factor analysis (CFA) and reliability and validity testing. The US sample served to assess the scale's robustness across a distinct, but similar, population, consistent with prior methodological approaches in this area. Sample sizes of approximately 500 participants in each cohort were selected based on two

considerations. First, best practices suggest including approximately 10 participants per item in EFA and CFA;[18] with a total of 40 candidate items for inclusion, we conservatively estimated 500 participants per cohort. Second, our aim was to use demographically representative samples to improve our generalizability, and including a 500-subject sample size in each population permitted representative demographic stratification. Psychometric term definitions are included in the Appendix.

To maximize the generalizability of our scale to all current and future vaccine delivery technologies, we chose to benchmark all responses against respondents' feelings about vaccine delivery with a traditional needle and syringe. Responses to all items in the scale were therefore scored on a 5-point Likert scale ranging from 'much less than with a needle and syringe' to 'much more than with a needle and syringe', with a neutral option 'the same as with a needle and syringe' included. Items were reverse coded, as needed, to yield a final scale in which higher scores represent a greater degree of support for the delivery technology under question than with a needle and syringe.

This research was approved by the University of Oxford Medical Sciences Interdivisional Research Ethics Committee (approval reference R81585/ RE001). Participants were reimbursed less than £5 for their time. All participants provided informed consent for participation and were permitted to withdraw at any time. Demographic details were self-declared and linked to responses using a unique 25-character alphanumeric identifier. Demographic data included age, self-identified gender, self-identified race and ethnicity, employment status, educational background, household income, and COVID-19 vaccination status.

Qualitative analyses were performed using NVivo release 1.4 (QSR International, Burlington, Massachusetts, USA) and statistical analyses were performed using Stata for Mac version 16 (Stata Corporation, College Station, Texas, USA). A p-value of 0.05 was considered significant for all analyses.

5.2.2 Domain and Item Generation

To develop a set of domains for consideration, we relied on a three-part approach. First, a literature review was performed to determine what domains should be considered based on historical precedent. Second, the two 50-member survey cohorts were used to expand the range of domains for inclusion by asking both open- and closed-answer questions regarding priorities for vaccine delivery technologies. Finally, this qualitative research was followed by iterative feedback from the Oxford Vaccine Group and the Engineering Science biology/ in vivo cluster meetings.

Closed-ended responses were reported with a 5-point Likert scale while open-ended responses were recorded in free text format. Open-ended questions included those such as: ‘What are your priorities when deciding on how much you value a vaccine delivery technology?’ A draft survey was also provided, with respondents given the opportunity to answer whether any items should be added, modified, or deleted, and whether the answer choices and questions were clear and appropriate (Boateng et al, 2018; Kantor & Kantor, 2021; Kantor et al, 2023; Watkins, 2022).

5.2.3 Domain and Item Reduction and Exploratory Factor Analysis

Item reduction was performed by reassessing domain relevance, deleting redundant items, examining polychoric (item-item) and polyserial (item-test) correlations, and deleting items with low communality or factor loadings (Cappelleri et al, 2014). Both Bartlett's test of sphericity and the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy were assessed to confirm that data were amenable to factor analysis (Lorenzo-Seva & Ferrando, 2021). A combination of an iterative approach, parallel analysis, scree plot, and assessment of the minimum average partial correlation were used to assess the number of domains to retain for inclusion (Dinno, 2009; Horn, 1965; Watkins, 2022). EFA with categorical variables was performed using a polychoric correlation matrix with iterated principal factor analysis followed by promax oblique factor rotation (Flora & Curran, 2004; Holgado–Tello et al, 2010).

5.2.4 Confirmatory Factor Analysis

We performed confirmatory maximum likelihood factor analysis with structural equation modelling and a Satorra-Bentler scaled test statistic on the US sample using the model developed in the UK sample to confirm the robustness of our model. Goodness of fit indices were assessed, including the root mean squared error of approximation (RMSEA), the standardized root mean squared residual (SRMR), and the comparative fit index (CFI). Values of RMSEA≤0.08, SRMR≤0.08, and CFI≥0.90 suggest acceptable fit, while values of RMSEA≤0.05, SRMR≤0.06, and CFI≥0.95 suggest excellent fit (Browne & Cudeck, 1992).

5.2.5 Reliability and Validity

Reliability as internal consistency was assessed using Cronbach's alpha for the overall OBSRVT scale and each subscale (Streiner, 2003).

We assessed convergent validity using the Oxford Needle Experience (ONE) scale, a validated scale that measures public attitudes to needles (Kantor et al, 2023). Discriminant validity was assessed using a single question regarding political outlook ('I consider myself to be a conservative person'). For both convergent and discriminant validity, the Pearson correlation coefficient was used.

5.3 Results

5.3.1 Characteristics of the Representative Samples

5.3.1.1 UK EFA Sample

A total of 541 individuals initially accessed our survey; of these, 21 failed to complete the survey (yielding a completion rate of 96.1%) and 20 (3.8%) failed at least one attention check, yielding 500 unique participants for analysis. The median (interquartile range [IQR]) age was 48 (33, 59) and respondents ranged in age from 18 to 82. Overall, 52.3% of respondents identified as female and the median (IQR) household income was £45,000 (26,000, 65,000). Additionally, 15.6% of the participants identified as non-white, 42.3% reported having no children, 47.3% reported working full-time, and 91.4% reported receiving at least one vaccination for COVID-19.

5.3.1.2 US CFA Sample

A total of 527 individuals initially accessed our survey; of these, 24 failed to complete the survey (yielding a completion rate of 95.4%) and 22 (4.4%) failed at least one attention check, yielding 501 unique participants for analysis. The median (interquartile range [IQR]) age was 47 (32, 59) and respondents ranged in age from 18 to 95. Overall, 50% of respondents identified as female and the median (IQR) household income was $65,000 (40,000, 100,000). Additionally, 34.3% of the participants identified as non-white, 44.5% reported having no children, 56.7% reported working full-time, and 81.4% reported receiving at least one vaccination for COVID-19.

5.3.2 Domain and Item Reduction and Exploratory Factor Analysis

Our iterative approach that included a literature review, open-ended questions from pilot samples (including a vaccine hesitant sample), and expert feedback from the Oxford Vaccine Group and Engineering Science biology cluster, yielded 10 potential domains including effectiveness, safety (including risks and side effects), ease of use (convenience), acceptance, comfort, cultural acceptability, likelihood of recommendation to others, perceived advantages, perceived disadvantages, and concerns (including anxiety or fear) associated with the delivery approach. We explored a total of 40 items for inclusion in the scale; 25 items were ultimately culled due to redundancy, low item correlations, and low factor loadings, yielding a final 15-item scale.

Bartlett's test of sphericity ($p<0.0001$) suggested the data were appropriate for EFA, which was confirmed by the KMO measure of sampling adequacy, which was excellent at 0.95. A scree plot, minimum average partial correlation analysis, and parallel analysis suggested a two to four factor solution (Dinno, 2009; Garrido et al, 2011; Ledesma et al, 2015).

5.3.3 Domains Included in the Final Scale

The four domains ultimately included in the scale were acceptance (6 items), effectiveness (4 items), confidence (3 items), and concerns (2 items). The acceptance domain reflects a generally positive response to the vaccine delivery technology in question, including likelihood of vaccination, willingness, preference, and comfort. Effectiveness, in contrast, reflects the degree to which the respondent believes that the delivery technology will be effective in vaccine delivery, providing immunity, and protection. Confidence highlights the perceived ease of use, ability of healthcare professionals to deliver the vaccine, and perception of vaccine safety. Finally, the concerns domain addresses fear or worry regarding side effects or safety associated with vaccine delivery using the technology under study.

Factor loadings are included in Table 5.1 and the OBSRVT scale is presented in Table 5.2.

Item	Acceptance	Effectiveness	Confidence	Concerns
I would be likely to get vaccinated if this vaccine delivery method were available.	**0.9377**	-0.0221	-0.0817	-0.0374
I am willing to try this vaccine delivery method in the future.	**0.8894**	-0.0169	0.0593	-0.0194
I would prefer this vaccine delivery method.	**0.8703**	-0.0124	0.0634	0.0175
I believe this vaccine delivery method should be widely adopted.	**0.7570**	0.2215	-0.0222	-0.0429
I would recommend this vaccine delivery method to friends and family.	**0.7560**	0.1965	0.0220	-0.0286
This vaccine delivery method is comfortable.	**0.6434**	-0.2154	0.0945	0.0015
I trust that this delivery method will provide immunity.	0.0166	**0.9844**	-0.0676	-0.0112
I am confident in the long-term effectiveness of this vaccine delivery method.	0.0428	**0.9013**	-0.0279	0.0213
I believe this vaccine delivery method would be effective in preventing disease.	-0.0230	**0.8884**	0.0284	-0.0126
I believe this method would be effective in delivering the vaccine as needed.	0.0532	**0.7948**	0.1292	0.0155
This vaccine delivery method can be easily used for the general population.	0.1562	-0.0909	**0.7464**	0.0712
I am confident that health professionals can safely deliver vaccines using this method.	0.0431	0.1287	**0.7301**	0.0389
I feel this vaccine delivery method is safe.	0.0623	0.0129	**0.4538**	-0.2355
I am fearful of the potential side effects of this vaccine delivery method.	-0.0516	0.0352	0.0740	**0.8937**
I worry about the safety of this vaccine delivery method.	-0.0497	-0.0307	-0.0155	**0.8479**

Table 5.1 Rotated factor loadings for the OBSRVT.

Factor loadings >0.40 (our a priori minimum for consideration) are bolded.

#	Item	Domain
1	I would be likely to get vaccinated if this vaccine delivery method were available.	Acceptance
2	I am willing to try this vaccine delivery method in the future.	Acceptance
3	I would prefer this vaccine delivery method.	Acceptance
4	I believe this vaccine delivery method should be widely adopted.	Acceptance
5	I would recommend this vaccine delivery method to friends and family.	Acceptance
6	This vaccine delivery method is comfortable.	Acceptance
7	I trust that this delivery method will provide immunity.	Effectiveness
8	I am confident in the long-term effectiveness of this vaccine delivery method.	Effectiveness
9	I believe this vaccine delivery method would be effective in preventing disease.	Effectiveness
10	I believe this method would be effective in delivering the vaccine as needed.	Effectiveness
11	This vaccine delivery method can be easily used for the general population.	Confidence
12	I am confident that health professionals can safely deliver vaccines using this method.	Confidence
13	I feel this vaccine delivery method is safe.	Confidence
14	I am fearful of the potential side effects of this vaccine delivery method.	Concerns
15	I worry about the safety of this vaccine delivery method.	Concerns

Table 5.2 The OBSRVT scale.

All items are scored using a 5-point Likert scale ranging from much less than with a needle and syringe to much more than with a needle and syringe. Higher scores thus reflect a greater degree of desirability relative to traditional needle and syringe-based vaccination, with a score of 45 suggesting equivalence to needle and syringe-based vaccination. Items 14 and 15 are reverse scored. The total score can range from 15 to 75.

5.3.4 Confirmatory Factor Analysis

Using the final four-factor 15-item model and structural equation modelling with a Satorra-Bentler scaled test statistic on the US sample, we demonstrated very good goodness of fit characteristics, with RMSEA=0.057, SRMR=0.053, and CFI=0.958.

5.3.5 Scale Reliability and Validity

Reliability as internal consistency was measured using Cronbach's alpha; values for the overall scale were very good at 0.92. Values for each of the four subscales are included in Table 5.3.

	UK Sample	**US Sample**
Acceptance	0.91	0.92
Effectiveness	0.92	0.93
Confidence	0.70	0.71
Concerns	0.84	0.81
Overall OBSRVT	**0.91**	**0.92**

Table 5.3 Subscale and overall reliability (assessed using Cronbach's alpha).

Convergent validity between the OBSRVT score and the Oxford Needle Experience (ONE) scale score was 0.31 ($p<0.0001$), and discriminant validity between the OBSRVT scale and a single unrelated question ('I consider myself to be a conservative person') was -0.03 ($p<0.0001$), supporting our validity argument.

5.4 Discussion

This manuscript outlines the initial development and validation of the OBSRVT, the first scale to assess the public's attitude to novel vaccine delivery technologies. We developed a concise, feasible 15-item scale over four domains: acceptance, effectiveness, confidence, and concerns. By benchmarking responses relative to needle and syringe-based vaccination, we aimed to maximize the utility of the scale in assessing both existing and future vaccine delivery technologies. This may represent an important next step in better understanding public priorities regarding novel vaccine delivery technologies.

Indeed, prior studies evaluating a range of delivery technologies from microneedle array patch approaches to nasal sprays routinely report outcomes such as efficacy, safety, and stability (Baker et al, 2023; Forster & Junger, 2022; Kumar et al, 2022b; Nguyen et al, 2021b), but acceptance is generally reported without a validated or standardized instrument (Ma et al, 2023b; Qu et al, 2024). Given the importance of public buy-in, both to decrease hesitancy and justify the significant public and private cost of this research, this gap is striking and highlights the potential value of the present work.

We adopted a benchmarking approach to rate novel vaccine delivery technologies relative to the ubiquitous needle and syringe-based vaccination approach. Benchmarking not only allows us to assess preferences relative to needle and syringe vaccines, but also to compare different alternative vaccine delivery technologies to each other, since each is rated relative to a stable, familiar, and immutable benchmark.

Each of the four domains in our scale—acceptance, effectiveness, confidence, and concerns—includes a different number of items. Since the scale is scored on a Likert scale ranging from one ('much less than with a needle and syringe') to five ('much more than with a needle and syringe'), a score of three on each item suggests equivalence to needle and syringe ('the same as with a needle and syringe'). Thus, the neutral cut-offs for each subscale, as outlined in Table 2, are 18, 12, 9, and 6 (yielding an overall scale cut-off of 45), which allows for easy benchmarking of the technology under investigation; values higher than the cut-offs suggest that the technology under investigation is preferred to a needle and syringe. This approach is particularly powerful in the dynamic environment of vaccine delivery technology development, as the OBSRVT is flexible and adaptable enough to be used to evaluate future technologies that have not yet been considered.

5.4.1 Study strengths

The benchmarking approach is also powerful on a practical level; sample data (Table 5.4) for an assessment of a theoretical nasal spray vaccine, for example, suggest that the overall feelings regarding this approach in the UK and US are similar, with the same median score in each population, bolstering our validity argument. Even these sample data could suggest to researchers, funders, and governments that the public largely sees nasal spray as superior to needle and syringe-based vaccination, but that concerns regarding effectiveness (with scores slightly below the threshold, suggesting concern relative to needle and syringe-based vaccination) would need to be addressed carefully to encourage broad adoption. These findings also

further support the face validity of the OBSRVT scale, since (as expected) participants were enthusiastic about the potential for nasal vaccine delivery based on acceptance scores while—not surprisingly—they had some residual concern regarding effectiveness, presumably given the relative novelty of the technology.

	UK Sample Median (IQR)	**US Sample Median (IQR)**	**Cut-offs for needle and syringe equivalence**	**p-value**
Acceptance	21 (18-25)	21 (17-24)	18	0.02
Effectiveness	10 (8-12)	11 (8-12)	12	0.07
Confidence	10 (9-12)	10 (9-11)	9	0.02
Concerns	6 (6-7)	6 (4-7)	6	0.27
OBSRVT Overall	48 (43-54)	48 (41-54)	45	0.17

Table 5.4 Representative comparison of scores for the OBSRVT scale when assessing a nasal spray vaccine between UK and US Samples.

The table presents the median and interquartile range (IQR) values for responses in the domains of acceptance, effectiveness, confidence, and concerns, as well as the overall OBSRVT score. P-values are derived from the Mann-Whitney U test, assessing the statistical significance of differences in the distributions between the two independent samples. Cutoffs for needle and syringe-based equivalence are based on a score of three for each Likert scale item; this is computed easily based on the number of items in each domain. A p-value less than 0.05 suggests a statistically significant difference.

Our work has several other strengths. First, we used a rigorous iterative process for domain and item development, including qualitative data from open-ended questions, to ensure that we could capture public priorities adequately. Second, we specifically included a vaccine hesitant cohort when initially soliciting feedback to increase the chances that we would capture priorities of a traditionally marginalized—yet critical for increasing novel vaccine adoption—group. Third, we used large demographically representative samples for our exploratory and confirmatory factor analyses to improve the generalizability of our findings. Fourth, to

assess the robustness of our model, we used a UK sample for exploratory factor analysis and a US sample for confirmatory factor analysis. Fifth, we performed our EFA using polychoric correlation matrices, an approach designed specifically for categorical data, to ensure that the factor analysis was performed appropriately given the Likert-scale responses. Finally, as alluded to above, the face validity of the scale was supported by the similarity in the baseline scores between the UK and US populations as well as the lower scores seen for the effectiveness subscale.

By design, there is some overlap between the items included within each of the scale's domains. For example, in the Confidence domain, item 12 ('I am confident that health professionals can safely deliver vaccines using this method') addresses confidence in the perceived ability of health professionals to *practically use* the delivery technology in question, while item 13 ('I feel this vaccine delivery method is safe') addresses a more general sense of confidence in the *overall perceived safety* of the technology itself. The distinction between these items may be particularly notable when exploring novel vaccine delivery technologies that may be perceived as technically difficult to deliver in a reproducible and consistent manner.

5.4.2 Study limitations

Limitations of our study include our use of a survey panel for subject recruitment, so that participants may not truly reflect the population of interest, though this was mitigated by relying on large demographically representative samples. Indeed, the essentially identical scores seen in the UK and US populations (Table 5.4) suggests that our findings are indeed robust and thus likely representative.

Second, our domains include an uneven number of items, and only two items were included in the concerns domain; still, since those items address the two key area of potential concern for participants—side effects and safety—and since their loading and structure is statistically valid, the practical utility of the factorial structure of our scale is supported. Third, while we included representative samples from both the UK and US in this study, future studies exploring whether our scale is valid across disparate cultural contexts, as well as invariance testing and test-retest reliability assessments, may be warranted. Fourth, most of the items in our scale were framed to elicit acceptance; while this is in part a function of our domain choices, and we included two items with reverse scoring in the Concerns domain, it is possible that this approach could introduce bias. Fifth, the class imbalance inherent in our outcome may affect the interpretation of operating characteristics. While AUC is a rank-based metric that does not depend mechanically on class proportions, modest positive-class sizes can theoretically increase variance in sensitivity estimates at individual thresholds. This is, however, partially mitigated by our use of stratified bootstrap confidence intervals, which preserved the outcome ratio within each replicate, and by the consistency of AUC estimates across five independent samples with varying prevalence. Finally, given the potential for baseline attitudes to needle and syringe-based vaccination to change over time, individual ratings may not remain stable. Thus, future studies assessing the responsiveness of the OBSRVT and its subscales—that is, how individual scores change over time as a result of the evolution of preferences, changes in circumstance, or a history of exposure—may be considered.

5.4.3 Conclusions

This manuscript presents data on the development and validation of the OBSRVT, the first scale designed specifically to assess the public's attitudes to next-generation vaccine delivery technologies. Given the public and private outlay of research funding in these areas (Gouglas et al, 2018; Lalani et al, 2023), and the potential for such technologies to fundamentally shift the ways in which vaccine delivery is effectuated, a formal validated scale is necessary to better assess public buy-in, better justify research funding, better aid in de-risking technologies, and better understand potential acceptance bottlenecks prior to full vaccine technology deployment. By developing a short yet comprehensive 15-item, four-domain scale that captures key considerations for vaccine delivery technologies—acceptance, effectiveness, confidence, and concerns—we hope that our work may serve as a tool for investigators as vaccine delivery technologies continue to evolve.

Chapter 6: Public attitudes towards intranasal and ultrasound-mediated vaccine delivery: A cross-sectional study in the United Kingdom and United States

6.1 Introduction

Next-generation vaccine delivery technologies offer significant advantages over needle and syringe-based delivery, including mitigating cold-chain requirements (Menon et al, 2021), reduced vaccine dosing needs (Schnyder et al, 2021), avoidance of adjuvants (Hettinga & Carlisle, 2020), avoidance of sharps injuries (Farag et al, 2018; Hacker et al, 2023), and lowered barriers for vaccine administration staff requirements (Cahn et al, 2023; O'Shea et al, 2021; Williams, 2013). Vaccine hesitancy is a major public health challenge, and while its aetiology is multifactorial (Fenta et al, 2023; Gauna et al, 2025), our work has previously suggested that needle fear may represent a risk factor for refusal (Freeman et al, 2021). Thus, next-generation vaccine delivery technologies are being developed both to potentially increase rates of vaccination and improve vaccine delivery logistics. Moreover, by avoiding the need for sharps disposal, or even medical waste disposal, alternative delivery technologies such as nasal spray and ultrasound could lead to considerable cost savings as well (Hagedorn et al, 2025).

Despite a significant and growing body of literature exploring the technical feasibility of alternative vaccine delivery technologies, little granular detail is known regarding the nature of public preferences for such approaches and the degree to which these preferences are associated with vaccine hesitancy. One group in Italy noted that those who are vaccine hesitant prefer nasal spray as a delivery technology

over needle and syringe-based vaccination (Boragno et al, 2023). Another study in Hong Kong found that both vaccine hesitancy (as a general phenomenon) and preference for nasal spray delivery increased after the COVID-19 pandemic, and that preference for nasal spray vaccine was associated with increasing vaccine hesitancy, suggesting that those who are vaccine hesitant, in particular, prefer non-needle approaches to vaccination (Yuan et al, 2024). A study in the US state of Georgia found that 29% of respondents would be more likely to vaccinate if offered alternative delivery methods rather than needle and syringe, and that those with low levels of vaccine hesitancy were significantly (adjusted prevalence ratio 1.40) less likely to be impacted by vaccine delivery technology—highlighting the possible importance of vaccine delivery technology in reaching the vaccine hesitant population (Melot et al, 2025).

More generally, there appears to be a preference for nasal spray over needle and syringe delivery, and indeed the prevalence of needle fear is higher in those who are under-vaccinated (Taddio et al, 2022). A study in China found that 59% of respondents favoured nasal spray over needle and syringe-based on stated preferences (Ma et al, 2023a). Another study evaluating preferences of children for seasonal influenza vaccines similarly found 69% preferred nasal spray delivery using conjoint analysis (Flood et al, 2011). Studies evaluating other alternative vaccine delivery technologies in clinical trial settings have similarly found a preference for non-needle alternatives (Fernando et al, 2018; Frew et al, 2020; Griffin et al, 2017; Norman et al, 2014).

While several studies have demonstrated preferences for alternative delivery technologies over needle and syringe, this has generally been assessed in the context of pre-rollout marketing or clinical trials, often using single unvalidated preference questions. No previous studies have used validated instruments to assess these preferences, and none have addressed preferences for ultrasound-mediated vaccination. This is particularly important as novel vaccine delivery technologies may be critical for reaching those who are vaccine hesitant.

We therefore sought to explore the preferences for two alternatives to needle and syringe-based vaccination approaches, nasal spray delivery—a model of a well-established and already commercialized delivery technology, evaluated as a baseline for understanding preferences to non-needle based vaccine delivery familiar to many respondents—and ultrasound-mediated transcutaneous delivery, a novel vaccine delivery technology that is currently under development (Hettinga & Carlisle, 2020; Hettinga et al, 2023). Rather than seeking to elicit the actual amounts individuals would be willing to pay for vaccination, our goal was simply to develop a quantitative approximation of relative preference across delivery technologies using multiple complementary approaches, including validated scales, stated preferences, willingness to pay, and a discrete choice experiment. In addition to avoiding the pain, potential needle-stick risk, disposal costs, and vaccinator training associated with needle and syringe-based vaccination, ultrasound-mediated vaccination into the dermal compartment provides intrinsic immune cell activation, potentially lowering the requirement for adjuvants, which may make it particularly appealing to the vaccine hesitant population (Hettinga & Carlisle, 2020; Tezel et al, 2005).

6.2 Methods

6.2.1 Population, Sample Size, and Framing

The population and data collection approach for the UK and US samples (~500 participants in each sample) was based on a methodology that we have previously employed in developing validated scales as detailed in the preceding chapters (Kantor & Kantor, 2021; Kantor et al, 2024a; Kantor et al, 2024b; Kantor et al, 2025; Kantor et al, 2023). We developed our surveys using Qualtrics (Qualtrics International, Provo, Utah, USA) and solicited responses using a survey panel approach through the academic survey site Prolific Academic (Oxford, UK). We initially performed a pilot study on a 50-respondent convenience sample to refine our scale items and approach and develop meaningful sample size estimates using both closed- and open-ended responses.

Sample size estimates were based on the primary outcome of interest, score on the Oxford Benchmark Scale for Rating Vaccine Technologies (OBSRVT) scale. This validated scale was developed specifically to assess respondents' enthusiasm for alternative vaccine delivery technologies relative to traditional needle and syringe delivery. Sample size calculations were performed based on pilot data; the mean OBSRVT score for nasal spray mediated vaccination was 46 with a standard deviation of 7.5, compared to the OBSRVT benchmark needle and syringe equivalence score of 45. With a power of 0.8 and alpha of 0.05, this yielded a necessary sample size of 444 for each population of interest.

We then used these sample size estimates to study a population in the UK, a demographically representative sample quota-matched by age, sex, and ethnicity (based on data from the UK Office of National Statistics) using a survey panel

approach. This was followed by sampling a second population in the US, quota-matched by the same demographic characteristics based on 2015 US national census bureau data. These populations were distinct from those analysed in Chapters two through five, as the Prolific Academic platform allowed us to exclude prior study participants based on their unique alphanumeric identifiers.

We framed ultrasound and nasal spray delivery using non-judgmental language. We specifically probed the pilot study group regarding our framing to ensure that the descriptions were clear. For ultrasound, we described the approach as: 'A gel containing a vaccine is applied to your skin. An ultrasound device is then held over the area for less than two minutes. The sound waves from the ultrasound help the vaccine move through the skin, allowing it to be absorbed'. For nasal spray, we based our framing on the description used in the prescribing information for an approved vaccine relying on nasal spray technology: 'A syringe containing the vaccine is inserted in one of your nostrils. The plunger is pressed, and the vaccine sprays up your nostril. This is then repeated on the other nostril'.

This work was approved by the University of Oxford Medical Sciences Interdivisional Research Ethics Committee (approval reference R81585/ RE001). All participants were reimbursed less than £5/$5 for their time, provided informed consent for participation, and were permitted to withdraw at any time. Demographic details were self-declared and linked to responses using a unique 25-character alphanumeric identifier. Baseline demographic data included age, gender, race and ethnicity, employment status, educational background, household income, and COVID-19 vaccination status. Multiple attention check items were included in the

survey instrument, and results were excluded from the analysis if a respondent failed two or more attention checks.

Qualitative analyses were performed using NVivo release 1.4 (QSR International, Burlington, Massachusetts, USA) and statistical analyses were performed using Stata for Mac version 16 (Stata Corporation, College Station, Texas, USA). A p-value of 0.05 was considered significant for all analyses. Psychometric term definitions are included in the Appendix.

6.2.2 Validated Scales

We used the OBSRVT as our primary outcome measure to assess relative preferences of participants for nasal spray and ultrasound-mediated vaccination over needle and syringe. OBSRVT scores range from 15 to 75, with 45 translating to equivalency to needle and syringe-based vaccination and higher scores suggesting a preference for the ultrasound or intranasal alternative.

Participants also completed both the Oxford Needle Experience (ONE) scale (Kantor et al, 2023) and the Oxford Vaccine Hesitancy Scale (OVHS) (Kantor et al, 2024b) for secondary analyses and stratification. The ONE scale assesses overall attitudes to needles, with scores ranging from 19 to 95, and higher values suggesting more concern or hesitancy regarding needles. The OVHS scores range from 13 to 65, with higher scores reflecting greater degrees of vaccine hesitancy.

6.2.3 Willingness to Pay and Stated Preferences

Stated willingness to pay (WTP) was used as a surrogate measure to assess preference. This was solicited by directly presenting participants with a delivery scenario, as outlined above, and asking 'How much would you be willing to pay (or have the NHS or another insurer pay) for these approaches to vaccine delivery?' In the US, this wording was modified by removing a reference to the NHS. Participants were then presented with the three options for vaccination: needle and syringe, nasal spray, and ultrasound, and presented with a continuous range of values to choose using a slider from £50/$50 to £300/$300. The default slider position was set to £100/$100 both for convenience and because that is approximately the standard charge for COVID-19 vaccination in a large UK pharmacy chain (£98.95) for those interested in paying privately for vaccines (Boots Pharmacy, 2025). A demand curve was created by modelling the relationship between demand (proportion willing to pay) and price, based on stated WTP.

A similar approach was used for stated preferences, where participants were asked 'How desirable do you think each option is for vaccination?' They were similarly presented with the three options for vaccination and could choose a value ranging from 0 to 100, with a default starting slider position set to 50. The WTP and stated preferences items were presented prior to the discrete choice experiment to minimize the risk of anchoring respondents to values used in the discrete choice experiment.

6.2.4 Discrete Choice Experiment

We performed two separate binary choice discrete choice experiments (DCEs) to assess preferences for nasal spray and ultrasound-mediated vaccination. Each DCE was comprised of four binary choice tasks comparing the baseline of needle and syringe-based vaccination to the alternative delivery intervention, either nasal spray or ultrasound. Participants were therefore presented with eight separate DCE tasks. In each choice task, participants were asked whether they preferred the benchmark needle and syringe or the alternative vaccine delivery technology under study at a specified price.

The baseline needle and syringe option was set at £100/$100 and the alternative was presented at four different price levels (£125/$125, £150/$150 £200/$200, or £250/$250) in random order. Price levels were chosen based on results from our pilot study. Participants were asked which option they would prefer at a stated price level.

6.2.5 Associations with Needle and Vaccine Hesitancy

We evaluated our primary outcomes—OBSRVT score, stated WTP, and stated preferences—by ONE and OVHS quartiles to better understand the way preferences for a particular vaccine delivery technology were associated with differential levels of needle or vaccine hesitancy. We hypothesized that increasing ONE and OVHS pain subscale quartiles would be associated with decreased enthusiasm for needle-based vaccination and increased valuation of alternative delivery technologies. ONE and

OVHS results were divided into quartiles, and primary outcome measures were assessed by quartile; Cuzick's test for trend was used to assess significance.

6.2.6 Statistical Analysis

Baseline continuous and categorical variables were compared using t-tests and chi-squared tests, respectively. For the DCE, we used conditional (fixed-effects) logistic regression to model the probability of choosing each alternative delivery technology over needle and syringe, conditioned on each participant's choice set. The primary independent variable was price. The model estimates a price coefficient (β_{price}) to demonstrate how changes in cost affect the log-odds of choosing the new intervention on a latent utility scale. This yields an equation $Utility_{ij}=\beta_{price}\times(Price_{ij})+\varepsilon_{ij}$ where $Price_{ij}$ is the price of alternative *j* in choice set *i*, and ε_{ij} is an error term. Thus, a negative and statistically significant β_{price} suggests that increasing price lowers the likelihood of choosing the intervention. For ease of interpretation, we also presented data as odds ratios per £25/$25 price increase using the formula $OR=e^{\beta\times25}$ to outline the expected reduction in odds of choosing a given option given a £25/$25 price increase. This approach is grounded in random utility theory, which assumes that each respondent assigns a latent utility (an unobserved measure of overall desirability) to each alternative, and selects the option with the higher utility (Lancsar & Louviere, 2008).

We assessed model fit using the pseudo R^2 (McFadden's R^2), comparing the log likelihood of the fitted model with that of a null (intercept-only) model. Groups in which all choices were identical (i.e., always choosing baseline or always choosing the

new intervention) were dropped automatically by the conditional logistic estimation as they provided no within-choice variation.

6.3 Results

6.3.1 Samples

The pilot sample included 50 participants, with a mean (SD) age of 38.7 (13.7), an age range of 18 to 67, and of whom 48% were female and 75% reported previous COVID-19 vaccination. Open-ended responses suggested that all scale items, including the framing of nasal spray and ultrasound-mediated vaccination, were easy to understand.

6.3.2 UK Sample

In the UK sample, 507 respondents initially accessed our survey, while seven declined to complete and two timed out their responses. Of the remaining 498 respondents who completed the survey, seven were dropped for failing attention checks, yielding 491 unique surveys for analysis. The mean age was 47 (95% confidence intervals [CI] 45.6, 48.3) with a range of 18 to 83, 50.9% were female, and the mean household income was £49,624. Additionally, 16.1% identified as non-white, 51.5% were married, and 49.6% were working full time.

6.3.3 US Sample

In the US sample, 540 respondents initially accessed our survey; 18 declined to complete and 14 timed out their responses. Of the remaining 508 respondents who

completed the survey, eight were dropped for failing attention checks yielding 500 unique surveys for analysis. The mean age was 45.8 (95% CI 44.4, 47.3) with a range of 18 to 85, 49.8% were female, and the mean household income was $92,900. Additionally, 26.7% identified as non-white, 39.7% were married, and 48.2% were working full time.

6.3.4 Outcome Measures

6.3.4.1 OBSRVT

In the UK, the mean OBSRVT score for nasal spray was 46.50 (95% CI 45.59, 47.41), p=0.0013 versus needle and syringe, and for ultrasound was 47.98 (47.07, 48.88), p<0.0001 versus needle and syringe. The OBSRVT scores for ultrasound and nasal spray differed significantly (p=0.0066). In the US, the mean OBSRVT score for nasal spray was 47.04 (95% CI 46.02, 48.07), p=0.0001 vs needle and syringe, and for ultrasound was 47.41 (46.35, 48.48), p<0.0001 vs needle and syringe. OBSRVT scores for nasal spray and ultrasound did not differ significantly (p=0.4349). OBSRVT subscale scores are included in Table 6.1.

Sample	Delivery Technology	Scale / Sub-scale	Mean	95 % CI
UK	Nasal spray	OBSRVT Overall	46.50	45.59, 47.41
		Acceptance	19.37	18.82, 19.93
		Effectiveness	11.03	10.78, 11.29
		Confidence	9.92	9.72, 10.11
		Concerns	6.16	6.03, 6.28
	Ultrasound	OBSRVT Overall	47.98	47.07, 48.88
		Acceptance	21.06	20.53, 21.60
		Effectiveness	11.06	10.79, 11.32
		Confidence	9.54	9.33, 9.75
		Concerns	6.31	6.17, 6.44
US	Nasal spray	OBSRVT Overall	47.04	46.02, 48.07
		Acceptance	19.64	19.04, 20.23
		Effectiveness	11.19	10.92, 11.46
		Confidence	9.98	9.76, 10.20
		Concerns	6.22	6.08, 6.37
	Ultrasound	OBSRVT Overall	47.41	46.35, 48.48
		Acceptance	20.39	19.79, 20.98
		Effectiveness	11.00	10.71, 11.29
		Confidence	9.64	9.40, 9.87
		Concerns	6.38	6.23, 6.53

Table 6.1 Overall and subscale OBSRVT scores for nasal spray and ultrasound-mediated vaccination.

Equivalence to needle and syringe is when overall OBSRVT score is 45.

6.3.4.2 Willingness to Pay

In the UK population, the mean stated WTP for needle and syringe-based vaccination was £94.18, increasing to £98.25 for nasal spray and £117.93 to ultrasound. In the US, mean stated WTP for needle and syringe-based vaccination was $94.40, increasing to $101.47 for nasal spray and $119.58 for ultrasound (Table 6.2).

Sample	Delivery Technology	Mean WTP (£ or $)	95 % CI
UK	Needle & syringe	94.18	91.03, 97.33
	Nasal spray	98.25	94.31, 102.20
	Ultrasound	117.93	113.48, 122.39
US	Needle & syringe	94.40	91.34, 97.46
	Nasal spray	101.47	97.29, 105.66
	Ultrasound	119.58	114.83, 124.32

Table 6.2 Willingness to pay for various delivery technologies, by sample.

A demand curve for WTP for nasal spray and ultrasound in the UK and US is presented in Figure 6.1.

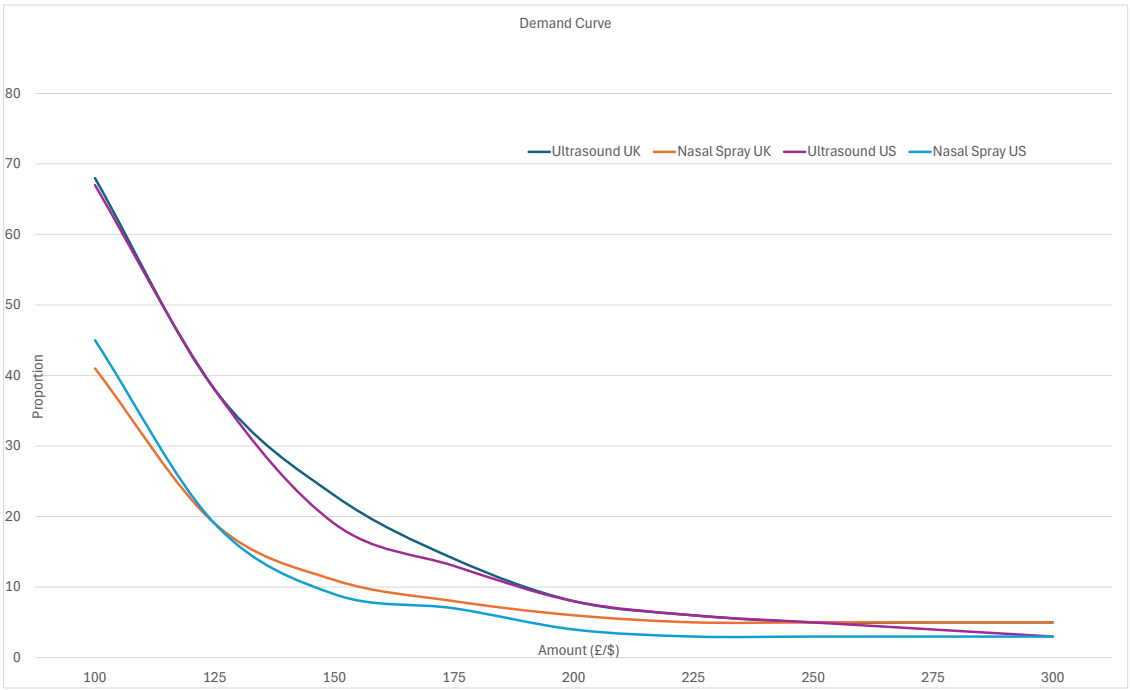


Figure 6.1 Demand curve illustrating the relationship between price and demand.

As a secondary analysis, we evaluated the association between respondents' WTP £100/$100 for each delivery technology and OBSRVT score. In the UK, the OR (95% CI) for ultrasound was 1.05 (1.03, 1.08), p<0.0001 suggesting that for every one-point increase in the OBSRVT score the WTP £100/$100 increased by 5%. This translated into a 71% increase in the odds of being willing to pay £100/$100 per

standard deviation increase in OBSRVT score. For nasal spray the OR (95% CI) was 1.03 (1.01, 1.05), p=0.001, translating to a 41% increase in odds of being willing to pay £100/$100 per standard deviation increase in OBSRVT score.

In the US, the OR (95% CI) was 1.07 (1.05, 1.09), p<0.0001 for ultrasound, translating to an 82% increase in odds per standard deviation. For nasal spray, it was 1.05 (1.03, 1.07), p<0.0001, translating to a 62% increase in odds of being willing to pay £100/$100 per standard deviation increase in OBSRVT.

6.3.4.3 Stated Preferences

In the UK, the desirability score (as defined in section 6.2.3) for needle and syringe-based vaccination was 48.38, 58.02 for nasal spray, and 67.56 for ultrasound. In the US, the mean desirability score for needle and syringe-based vaccination was 50.63, increasing to 57.87 for nasal spray and 65.29 for ultrasound (Table 6.3).

Sample	Delivery Technology	Mean	95 % CI
UK	Needle & syringe	48.38	45.70, 51.06
	Nasal spray	58.02	55.44, 60.59
	Ultrasound	67.56	65.10, 70.03
US	Needle & syringe	50.63	47.78, 53.48
	Nasal spray	57.87	55.19, 60.54
	Ultrasound	65.29	62.72, 67.87

Table 6.3 Stated preferences (0-100, with higher values representing more desirability) for vaccine delivery technologies by sample.

6.3.4.4 Discrete Choice Experiment

Conditional (fixed effects) logistic regression with price as the varying attribute in the UK demonstrated a statistically significant negative price sensitivity for both

delivery methods. For nasal spray, the coefficient was β = -0.0141266 with a McFadden R^2 of 0.2255 ($p<0.0001$). For ultrasound, the coefficient was β = -0.0046568 with a McFadden R^2 of 0.0333 ($p<0.0001$). Extrapolating to a £25 incremental price increase and converting the log-odds coefficients to odds ratios indicates that each £25 price increase reduces the odds of choosing the nasal spray option by 30% (OR=0.70) and ultrasound by 11% (OR=0.89). McFadden's R^2 suggests a greater price sensitivity for nasal spray over ultrasound; as price increases, the odds are higher that respondents will choose the reference option (needle and syringe-based vaccination) over nasal spray, while for ultrasound this effect is less pronounced.

In the US, the coefficient for nasal spray was β = -0.0203571 with a McFadden R^2 of 0.3572 ($p<0.0001$). For ultrasound, the coefficient was β = -0.017336 with a McFadden R^2 of 0.2964 ($p<0.0001$). This translates into a reduction in odds of choosing (per $25 price increase) the nasal spray option by 40% (OR=0.60) and ultrasound by 35% (OR=0.65). McFadden's R^2 suggests a slightly greater price sensitivity for nasal spray than ultrasound in this population.

6.3.5 Associations with Needle and Vaccine Hesitancy

In both the UK and US, OBSRVT scores for both ultrasound and nasal spray increased with increasing ONE scale quartile, suggesting that the relative preference for each vaccine delivery technology over needle and syringe-based vaccination was more pronounced for respondents who were more needle hesitant (Table 6.4).

	Delivery Technology	Q1 (Mean, SD)	Q2 (Mean, SD)	Q3 (Mean, SD)	Q4 (Mean, SD)
UK	Nasal spray	44.77 (9.42)	45.69 (9.85)	46.68 (10.28)	49.11 (10.95)
	Ultrasound	45.58 (9.11)	46.84 (9.82)	48.96 (10.04)	50.66 (10.89)
US	Nasal spray	43.21 (11.19)	45.86 (9.47)	47.67 (10.94)	51.81 (13.04)
	Ultrasound	42.46 (11.73)	46.20 (10.63)	47.66 (11.45)	53.49 (11.97)

Table 6.4 OBSRVT scores stratified by ONE scale quartile.

An OBSRVT score of 45 indicates equivalency with needle and syringe-based vaccination; higher scores reflect greater levels of acceptance. All p-values using Cuzick's non-parametric test for trend are p<0.0001.

A similar effect on OBSRVT scores was seen with each quartile increase in the OVHS pain subscale score in both the UK and US (UK ultrasound: 46.44 → 51.27; US ultrasound: 44.23 → 52.12), and for the OVHS deliberation subscale score and OVHS overall score in the US, as highlighted in Table 6.5.

Sample	Delivery Technology	OVHS Subscale	Q1 (Mean, SD)	Q2 (Mean, SD)	Q3 (Mean, SD)	Q4 (Mean, SD)	p-value
UK	Nasal Spray	Beliefs	45.97 (9.47)	46.68 (10.12)	45.55 (11.12)	47.42 (10.39)	0.375
		Pain	45.43 (9.58)	47.35 (10.23)	45.91 (10.41)	49.09 (11.08)	0.009
		Deliberation	47.05 (10.08)	46.08 (9.20)	45.82 (10.05)	47.21 (11.98)	0.757
		Overall	45.85 (9.10)	46.32 (10.02)	45.91 (11.33)	47.81 (10.71)	0.142
	Ultrasound	Beliefs	46.60 (10.07)	49.01 (10.47)	47.53 (10.29)	48.52 (9.48)	0.147
		Pain	46.44 (9.28)	49.52 (10.65)	47.45 (10.32)	51.27 (10.67)	<0.0001
		Deliberation	47.93 (9.57)	45.78 (10.08)	49.15 (9.89)	48.71 (11.48)	0.180

		Overall	46.71 (9.79)	47.86 (10.06)	48.04 (10.50)	49.60 (10.10)	0.004
US	Nasal Spray	Beliefs	46.21 (10.64)	48.44 (12.17)	48.73 (10.69)	45.69 (13.00)	0.763
		Pain	44.66 (11.13)	46.98 (9.97)	46.48 (11.99)	50.71 (13.28)	<0.0001
		Deliberation	45.68 (11.88)	47.16 (10.63)	45.85 (11.08)	49.26 (12.60)	0.030
		Overall	45.61 (10.19)	45.96 (11.56)	48.50 (10.16)	48.49 (14.17)	0.008
	Ultrasound	Beliefs	45.57 (12.11)	50.09 (10.43)	49.52 (11.41)	46.53 (13.34)	0.131
		Pain	44.23 (12.09)	47.50 (10.54)	46.73 (12.41)	52.12 (12.61)	<0.0001
		Deliberation	45.31 (11.43)	46.91 (11.90)	46.75 (12.32)	50.77 (12.37)	<0.0001
		Overall	44.64 (11.27)	47.06 (12.14)	49.32 (10.61)	49.60 (13.85)	<0.0001

Table 6.5 OBSRVT score by OVHS scale quartile.

An OBSRVT score of 45 indicates equivalency with needle and syringe; higher scores reflect greater levels of acceptance. p-values are for Cuzick's non-parametric test for trend.

WTP is similarly associated with increasing needle hesitancy; as ONE score increases, WTP for needle and syringe-based vaccination decreases (UK: 98.48 → 91.88; US: 101.87 → 89.47), while WTP for nasal spray increases (UK: 92.62 → 114.06; US: 96.61 → 117.48) and WTP for ultrasound-mediated vaccination increases dramatically (UK: 108.93 → 132.87; US: 102.66 → 142.88), as seen in Table 6.6.

	Delivery Technology	Q1 (Mean, SD)	Q2 (Mean, SD)	Q3 (Mean, SD)	Q4 (Mean, SD)	p-value
UK	Needle & Syringe	98.48 (33.54)	92.20 (28.00)	93.93 (27.84)	91.88 (37.76)	0.022
	Ultrasound	108.93 (43.03)	113.17 (48.18)	116.59 (46.01)	132.87 (53.51)	<0.0001
	Nasal Spray	92.62 (33.59)	89.60 (40.08)	96.95 (41.96)	114.06 (50.55)	0.001
US	Needle & Syringe	101.87 (37.98)	96.65 (29.90)	88.04 (24.17)	89.47 (34.53)	<0.0001
	Ultrasound	102.66 (42.76)	114.30 (48.66)	116.56 (48.43)	142.88 (61.41)	<0.0001
	Nasal Spray	96.61 (41.64)	92.20 (36.71)	99.39 (46.80)	117.48 (55.99)	0.001

Table 6.6 Mean willingness to pay (£/$) by ONE scale quartile.

p-values are for Cuzick's non-parametric test for trend.

WTP values also are associated with vaccine hesitancy: increasing OVHS quartile is associated with a decreased WTP for needle and syringe-based vaccination. Conversely, increasing OVHS pain subscale score is associated with an increased WTP for nasal spray, and a pronounced increased WTP for ultrasound (Table 6.7).

Sample	Delivery Technology	OVHS Sub-scale	Q1 (Mean, SD)	Q2 (Mean, SD)	Q3 (Mean, SD)	Q4 (Mean, SD)	p-value
UK	Needle & Syringe	Beliefs	97.63 (32.40)	97.63 (26.60)	93.17 (34.16)	87.36 (35.60)	0.002
		Pain	94.94 (26.01)	93.16 (37.07)	94.66 (32.66)	92.75 (39.91)	0.243
		Deliberation	94.72 (31.08)	96.24 (35.71)	93.61 (33.31)	91.84 (27.94)	0.589
		Overall	97.66 (26.71)	95.10 (32.60)	92.47 (33.37)	90.46 (36.66)	0.016

	Ultrasound	Beliefs	118.01 (43.54)	123.61 (48.19)	115.83 (48.11)	112.85 (56.46)	0.111
		Pain	110.50 (41.86)	121.37 (49.04)	119.51 (55.44)	132.04 (54.80)	0.001
		Deliberation	118.26 (46.66)	111.90 (45.42)	120.18 (50.60)	120.73 (55.43)	0.982
		Overall	115.17 (42.25)	118.98 (45.88)	121.39 (56.73)	118.00 (53.92)	0.835
	Nasal Spray	Beliefs	98.51 (37.96)	97.99 (38.58)	99.12 (45.39)	97.84 (51.49)	0.156
		Pain	89.72 (34.10)	96.84 (39.94)	100.47 (47.63)	117.96 (52.11)	<0.0001
		Deliberation	98.63 (44.25)	100.22 (42.52)	94.12 (42.22)	104.48 (42.37)	0.632
		Overall	94.86 (34.07)	95.67 (44.06)	97.95 (42.91)	105.29 (51.16)	0.400
US	Needle & Syringe	Beliefs	99.81 (34.53)	96.38 (34.53)	91.37 (24.20)	87.40 (33.15)	0.004
		Pain	98.66 (35.73)	92.66 (29.52)	88.03 (24.08)	92.51 (33.69)	0.074
		Deliberation	96.89 (32.62)	97.74 (37.28)	92.86 (30.52)	88.69 (27.26)	0.026
		Overall	98.19 (37.37)	99.77 (28.40)	93.21 (26.16)	84.85 (34.02)	0.001
	Ultrasound	Beliefs	117.67 (47.36)	126.23 (55.53)	122.88 (51.77)	115.09 (60.86)	0.389
		Pain	106.11 (43.11)	115.53 (51.25)	113.16 (47.74)	146.30 (60.96)	<0.0001
		Deliberation	113.24 (45.10)	118.59 (51.00)	120.88 (58.77)	126.57 (58.85)	0.178
		Overall	107.45 (40.11)	121.32 (50.28)	123.44 (50.53)	130.72 (69.75)	0.013
	Nasal Spray	Beliefs	105.05 (46.05)	100.55 (42.63)	100.28 (41.25)	97.84 (54.26)	0.043
		Pain	94.08 (42.20)	97.31 (42.12)	94.38 (47.47)	119.88 (53.33)	<0.0001
		Deliberation	103.89 (43.16)	102.44 (43.79)	97.90 (52.37)	100.32 (49.31)	0.168

	Overall	96.84 (37.35)	103.88 (46.00)	104.14 (47.42)	102.69 (56.65)	0.914

Table 6.7 Willingness to pay (in £ or $, as appropriate) by OVHS quartile.

p-values are for Cuzick's test for trend.

Stated preferences are similarly associated with needle hesitancy; as ONE scale score increases, preference for needle and syringe-based vaccination decreases (Table 6.8). No association was seen between increasing ONE scale quartile and nasal spray vaccination, though the preference for ultrasound as a delivery method increased with increasing needle hesitancy.

	Delivery Technology	**Q1 (Mean, SD)**	**Q2 (Mean, SD)**	**Q3 (Mean, SD)**	**Q4 (Mean, SD)**	**p-value**
UK	Needle & Syringe	58.09 (30.76)	49.54 (27.38)	44.88 (28.03)	39.73 (29.32)	<0.0001
	Nasal Spray	58.79 (30.36)	56.64 (29.67)	57.36 (28.48)	59.99 (26.73)	0.975
	Ultrasound	63.00 (28.74)	66.54 (28.32)	70.77 (25.79)	70.76 (26.00)	0.018
US	Needle & Syringe	66.04 (27.96)	60.73 (27.44)	42.60 (30.31)	32.26 (30.43)	<0.0001
	Nasal Spray	56.57 (33.73)	55.48 (29.92)	58.52 (28.52)	60.00 (28.74)	0.422
	Ultrasound	60.04 (31.14)	63.64 (28.85)	64.91 (28.26)	72.51 (28.19)	<0.0001

Table 6.8 Stated preferences (0-100, with higher values representing more desirability) by ONE scale quartile.

p-values are for Cuzick's test for trend.

Evaluating stated preferences by OVHS quartile revealed a significant decrease in preferences for needle and syringe-based vaccination as vaccine hesitancy quartile increases (Table 6.9). For nasal spray, no association was seen. For ultrasound-mediated vaccination, there was a significant increase in preference for this approach as OVHS pain and deliberation subscale quartiles increased in the US sample (p=0.003 and p=0.031, respectively).

Sample	Delivery Technology	OVHS Sub-scale	Q1 (Mean, SD)	Q2 (Mean, SD)	Q3 (Mean, SD)	Q4 (Mean, SD)	p-value
UK	Needle & Syringe	Beliefs	53.80 (29.11)	48.83 (30.45)	52.16 (27.95)	38.58 (27.86)	<0.0001
		Pain	53.38 (29.42)	46.16 (27.71)	49.33 (29.71)	37.35 (28.61)	<0.0001
		Deliberation	48.63 (29.59)	49.19 (28.81)	47.90 (29.88)	47.76 (30.35)	0.804
		Overall	51.54 (29.73)	53.61 (28.96)	48.60 (29.05)	39.13 (28.30)	0.001
	Nasal Spray	Beliefs	61.46 (27.20)	54.56 (29.78)	60.71 (27.72)	56.80 (29.48)	0.451
		Pain	59.42 (28.76)	56.84 (27.36)	56.50 (30.32)	58.09 (28.05)	0.477
		Deliberation	58.96 (27.43)	64.41 (25.78)	54.79 (29.51)	55.50 (32.15)	0.532
		Overall	59.09 (27.75)	59.25 (29.87)	57.83 (26.68)	56.12 (30.57)	0.407
	Ultrasound	Beliefs	70.12 (27.50)	69.09 (26.01)	64.49 (27.01)	65.45 (28.50)	0.060
		Pain	66.27 (27.53)	69.83 (25.59)	66.42 (29.57)	70.79 (25.00)	0.274
		Deliberation	67.65 (26.10)	63.50 (28.08)	69.17 (27.20)	69.86 (28.51)	0.169
		Overall	66.46 (27.96)	69.81 (26.13)	67.27 (27.48)	67.34 (27.33)	0.991
US	Needle & Syringe	Beliefs	58.75 (29.72)	46.57 (30.34)	49.02 (32.06)	41.88 (33.61)	<0.0001
		Pain	62.84 (29.06)	47.53 (31.38)	57.08 (27.82)	34.81 (30.57)	<0.0001
		Deliberation	53.54 (30.76)	52.04 (30.96)	55.42 (32.03)	42.26 (33.14)	0.013
		Overall	61.33 (28.41)	55.34 (30.33)	45.39 (31.41)	35.78 (32.33)	<0.0001
	Nasal Spray	Beliefs	60.21 (29.96)	57.51 (28.46)	57.31 (30.24)	54.47 (32.33)	0.161
		Pain	55.03 (32.65)	58.91 (29.42)	52.90 (27.73)	62.24 (28.57)	0.138

	Deliberation	59.09 (30.44)	59.36 (28.34)	55.49 (30.14)	56.54 (32.82)	0.532
	Overall	59.97 (30.08)	56.97 (30.21)	58.13 (28.59)	55.39 (32.60)	0.351
Ultrasound	Beliefs	67.09 (28.67)	69.80 (26.14)	63.04 (28.90)	60.55 (32.44)	0.076
	Pain	61.07 (30.01)	65.99 (28.54)	64.48 (28.06)	70.73 (29.19)	0.003
	Deliberation	62.74 (28.03)	66.11 (27.83)	63.03 (30.70)	68.85 (31.28)	0.031
	Overall	63.52 (28.13)	66.83 (29.32)	66.03 (28.13)	64.92 (32.22)	0.313

Table 6.9 Stated preferences (0-100, with higher values representing more enthusiasm) by OVHS quartile.

p-values are for Cuzick's test for trend.

6.4 Discussion

This manuscript outlines the use of multiple outcome measures—OBSRVT score, stated WTP, stated preferences, and a DCE—to assess the relative valuation of novel vaccine delivery technologies in demographically representative populations in the UK and US. Based on our primary outcome measure, OBSRVT score, both nasal spray and ultrasound as described are preferred over needle and syringe, with ultrasound yielding the highest scores. These results are echoed when using our other outcome measures, with stated WTP and stated preferences demonstrating a preference for ultrasound as the top choice of respondents. The stated WTP differential translated into a 25% average price premium that respondents were willing to pay for ultrasound over needle and syringe-based vaccination. This preference ranking was also supported by the DCE, where the lower price sensitivity for ultrasound in both the UK and US populations—and the smaller beta coefficient—

further supports that ultrasound is preferred over both needle and syringe and nasal spray as a delivery technology.

We also found that two related barriers to widespread vaccination—vaccine and needle hesitancy—may be mitigated by alternative next-generation vaccine delivery technologies. Indeed, OBSRVT scores increased per quartile of both the OVHS pain subscale and ONE scale for both nasal spray and ultrasound-mediated vaccination, though these increases were more pronounced for the latter. This shift in preferences among vaccine hesitant respondents towards nasal spray, and even more so towards ultrasound, was echoed using other outcome measures including stated WTP and stated preferences; as expected, stated preferences for needle and syringe-based vaccination decreased with greater levels of vaccine and needle hesitancy. Moreover, absolute preferences for needle-free technologies generally *increased* with both increasing vaccine and needle hesitancy. The magnitude of the preference for ultrasound was greater than nasal spray across all measures; this may be related to a higher perceived invasiveness of nasal spray over ultrasound, though further research is needed. Regardless, these findings highlight that while alternative delivery technologies are generally favoured over needle and syringe, this preference is even more pronounced for those who are needle or vaccine hesitant.

We used a DCE as one of the approaches to eliciting preferences and trade-offs in our study. Prior work has suggested that DCEs, with their ability to draw conclusions from multiple comparisons, are a reasonable approach to evaluating vaccine preferences, whether evaluating influenza vaccine preferences in China (Li et al, 2022) or attributes for a COVID-19 vaccine in the US (Kreps et al, 2020). Since the beta

coefficient reflects the slope of the price-utility curve, the preference for nasal spray was more price sensitive than the preference for ultrasound, and cost explains a greater proportion of the variance seen for nasal spray. Taken together, these findings highlight that while nasal spray is preferred over needle and syringe, the preference for ultrasound is even greater than nasal spray and is more resilient in the face of price increases, suggesting that it reflects a more durable preference.

The relative merits of ultrasound as measured by the OBSRVT scale were tempered somewhat by its novelty; in evaluating OBSRVT subscales, nasal spray, which is likely familiar to respondents given its availability in both the UK and US, demonstrated numerically higher scores on the confidence subscale in both samples (Table 6.1), though this difference was not formally tested. While speculative, this pattern could be consistent with a slight novelty penalty for ultrasound, so that as ultrasound is rolled out as an available and familiar technology its benefit over needle and syringe could potentially become even more pronounced. This echoes data in the COVID-19 pandemic context, where the novelty of mRNA vaccine technology modulated enthusiasm for this approach to vaccination in the general public (Leong et al, 2022). Similarly, in evaluating an intranasal COVID-19 vaccine in China, one study found that effectiveness and safety were the primary drivers of whether respondents would recommend the vaccine to others (Li et al, 2022). Finally, a study in the Gambia found that a history of familiarity with nasal spray was associated with an *increased* level of enthusiasm for it as a delivery option—in that study, 35% of respondents indicated a preference for nasal spray over needle and syringe, but this soared to 93% for those whose children had receive nasal spray vaccine (Armitage et al, 2018). This

underscores the importance of the novelty penalty and suggests that our findings may actually *underestimate* the relative preference for ultrasound-mediated vaccination, as increasing familiarity in the future may increase its already high level of acceptability even further.

6.4.1 Study strengths

Strengths of our study include our use of demographically representative populations in two different countries, enhancing the generalizability of our findings; the use of multiple validated scales; the use of multiple approaches to preference assessment, including stated preferences, WTP, OBSRVT score, and a DCE; stratification across OVHS and ONE scores; and robust statistical modelling. Indeed, with large demographically representative samples, we were powered to detect even subtle differences in the ways respondents rate technologies (Cohen's d of 0.133), and by using real-world price anchoring based on actual UK pharmacy charges our approach has greater external validity, potentially nudging observed elasticities towards representing market behaviour.

Importantly, however, our WTP values and DCE results were not designed to elicit the actual amounts that individuals would be willing to pay for vaccination; instead, these results were used to assess the *relative preferences* of respondents for one technology over another. Given that the cost and payor structure for vaccines is complex and variable across countries, vaccines, and individuals, our goal was simply to develop a quantitative approximation of preference to better understand public

preferences for different technologies and serve as a model for future work in this area.

The dose-response relationships in our findings also strongly support the validity arguments for the instruments used in these assessments: the OBSRVT, ONE scale, and the OVHS. The odds of being willing to pay £100/$100 for ultrasound increased by 71-82% for each standard deviation increase in OBSRVT and for nasal spray increased by 41-62% for each standard deviation OBSRVT increase, suggesting that increasing OBSRVT score indeed reflects increasing levels of preference. Similarly, the inverse association between increasing ONE score and WTP for needle and syringe-based vaccination supports the validity argument for the ONE scale. Finally, the association between increasing OVHS pain subscale scores and decreasing preference for needle and syringe-based vaccination (and increasing valuation of alternatives) bolsters the validity argument for this scale as well.

6.4.2 Study limitations

This study has important limitations. First, we only evaluated two alternative delivery technologies, nasal spray and ultrasound. Given that the purpose of this study was a) to assess preferences for ultrasound-mediated vaccination using nasal spray as a control and b) to develop an approach that can be deployed for other next-generation technologies in the future, this limitation is both expected and strategic given feasibility concerns. Indeed, this approach allowed us to establish proof-of-concept regarding ultrasound preferences and acceptability while also validating a transferable preference elicitation toolkit that can be used in future studies to de-risk

novel vaccine delivery technologies. Future studies evaluating other next-generation technologies may be helpful.

Second, our study relied on a text description of both nasal spray and ultrasound-mediated vaccine delivery, and therefore the results may be sensitive to framing or priming. We relied on non-judgmental language vetted for impartiality in our pilot study and based on published wording, when available, to mitigate this shortcoming. Still, future studies including a sensitivity analysis to framing and incorporating video of vaccine delivery approaches—or ideally live active or sham vaccine delivery using these technologies—may be helpful. Importantly, our findings reflect hypothetical stated preferences based on textual descriptions of vaccine delivery technologies. Participants had no actual experience with ultrasound-mediated vaccination, and therefore our measures capture anticipated preferences rather than post-experience evaluations. Our findings should therefore be interpreted as indicating which described attributes of vaccine delivery are most valued by respondents, with ultrasound-associated attributes consistently rated most favourably across multiple preference elicitation methods

Third, anchoring bias is always a risk when presenting price data; however, the baseline pricing ranges set on sliders was based on pricing at a public-facing website of a major UK pharmacy chain and we incorporated sliders with wide range of process to help minimize floor and ceiling effects. More importantly, while we present monetary values, our goal was relative valuation, rather than absolute valuation, and therefore anchoring—as long as it was unbiased—is less of a concern.

Fourth, the appropriateness of evaluating WTP in general for vaccines depends on perspective; we adopted a societal perspective as patients in both the UK and US do not generally pay directly for vaccines themselves on an out-of-pocket basis. This introduces potential for third-party payer bias, since respondents would generally be freer with other people's money.

Fifth, while we relied on validated scales, the scales were not tailored for each individual delivery technology (Berger et al, 2024). The OBSRVT, however, was validated specifically to compare next-generation vaccine delivery technologies to needle and syringe as a benchmark, though additional granular detail would likely be available if we used a specific validated scale for assessing acceptance of ultrasound-mediated vaccination, and such future studies are likely worthwhile.

Finally, we did not perform a cost-effectiveness analysis, and indeed recent literature suggests that next-generation vaccine delivery technologies could yield significant cost savings by dose-sparing and avoiding the need for sharps disposal (Hagedorn et al, 2025). Since other technologies, such as microarray patches, generally require medical waste disposal of the carrier or backing, shifting to an ultrasound-based delivery with single-use gel packets and reusable probes could result in even more profound cost savings, and future research modelling potential cost savings associated with ultrasound-mediated delivery may be warranted.

6.4.3 Conclusions

Using an array of outcome measures, including the OBSRVT scale, stated preferences, stated WTP, and a DCE, we found that ultrasound is preferred over both

nasal spray and needle and syringe as a vaccine delivery technology across demographically representative populations in the UK and US. As the first study to evaluate preferences for ultrasound-mediated vaccination and the first to stratify results of multiple alternative vaccination technologies by vaccine and needle hesitancy, this work hopefully can provide insight into the ways in which vaccine delivery technologies can solve both direct (cold chain, vaccine dosing, adjuvant use, logistics) and indirect (needle and vaccine hesitancy, uptake) vaccine delivery challenges. We also found that needle hesitancy, as measured by the ONE scale, and particularly vaccine hesitancy, as measured by the OVHS, is associated with these preferences, suggesting that vaccine hesitant patients are, perhaps paradoxically, potentially enthusiastic about next-generation vaccine delivery technologies, and that these technologies may therefore have the potential to mitigate vaccine refusal. Given that vaccine hesitancy may represent a more pressing barrier to vaccination than logistic challenges, these findings hold significant public health implications. Treating this work as a transferable toolkit for evaluating other vaccine technologies, and indeed other biomedical engineering challenges more generally, may hold significant promise for technology de-risking in the future across a range of applications.

Chapter 7: Conclusions: Developing the minimum desirable product by de-risking biotechnology ventures through early stage validated scale development and iterative human factors research

7.1 Human factors research

This thesis outlined the development and validation of a suite of instruments that may be used as part of the de-risking process for bringing alternative vaccine delivery technologies to market. Together, these instruments convert otherwise qualitative human factors data into actionable quantitative signals. By using these instruments in an iterative or layered fashion, researchers can gauge both the interest in and the potential impact of novel vaccine delivery technologies, ultimately lowering the bar to the development of feasible next generation delivery technologies by providing quantitative evidence of public acceptance that can strengthen funding applications, guide prototype iteration, and help de-risk investment decisions early in the product development cycle, as discussed further in Chapter 7..

Indeed, developing and implementing novel biotechnology solutions requires alignment between research and development (R&D) teams, investors and funding agencies, end users, and the public at large (Marx, 2024). Early-stage ventures traditionally focus on R&D and the race for a minimum viable product (MVP), shunting concern regarding human factors research—including an understanding of public preferences, use cases, demand, valuation, and likely adoption—downstream in the product development cycle.

Yet a significant proportion of biotechnology ventures fail due to lack of end-user buy-in or funding limitations, rather than technical or regulatory challenges (Paul et al,

2010), so that creating desirable products with a clear value proposition and viable odds of achieving product-market fit is critical at the earliest stages of development. Therefore, shifting human factors work earlier in the process and approaching it as an integral part of the R&D framework, rather than as an afterthought, and developing scales to quantify and understand otherwise unquantified human factors, may help reduce late-stage attrition rates (Alver et al, 2024).

This approach may allow nascent biotechnology ventures to leapfrog existing approaches with higher technology readiness levels by accelerating the creation of a product that is not only viable, but desirable as well (Figure 7.1). By satisfying quantitatively assessed thresholds for end-user and public acceptance, research groups can move from viability to desirability—as objectively measured by patient and public enthusiasm—and from the MVP to the minimum desirable product (MDP).

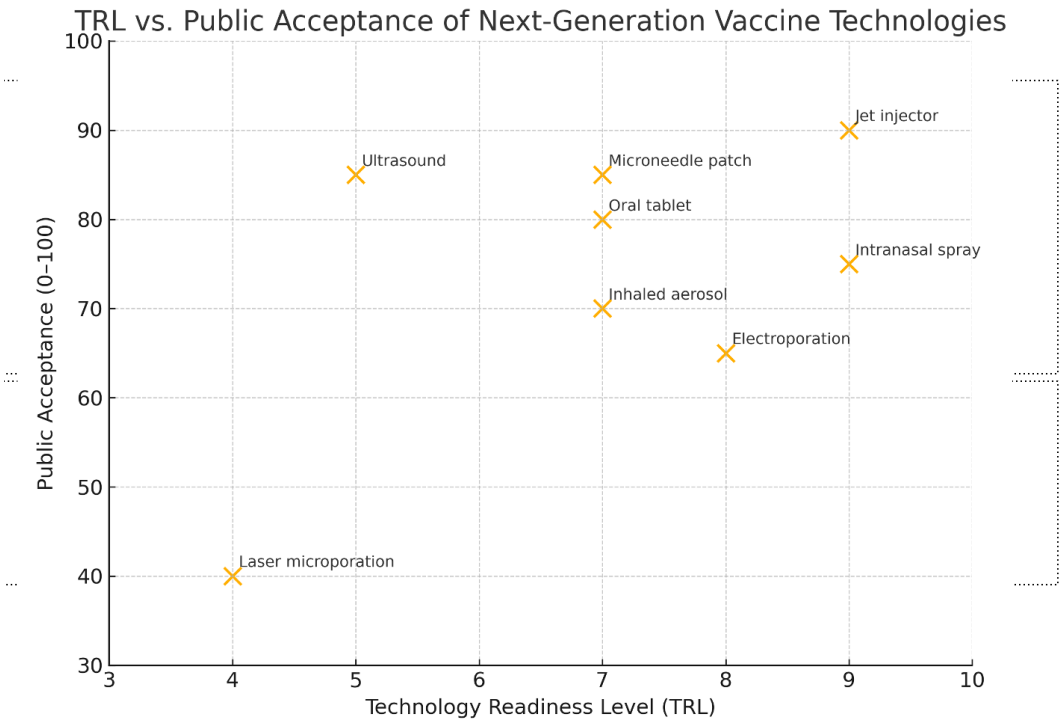


Figure 7.1 Technology readiness level (TRL) and public acceptance of vaccine delivery technologies. All values are demonstrative.

7.2 From uncertainty to risk: En-risk to de-risk

The need to quantify key outcomes in R&D is a given; yet in human-factors research, qualitative intuition often predominates, and the mere assertion that such research was conducted is sometimes seen—like a disclosure of conflicts of interest—to be all that is needed (van Tulleken et al, 2024). Instead, to make de-risking possible, quantifying human-factors uncertainty can convert uncontrollable unknown unknowns into manageable risks. To de-risk biotechnology, convert nebulous uncertainty to risk: en-risk in order to de-risk.

7.3 Instrument and Scale based risk quantification

Developing and validating novel instruments and scales is a replicable approach to rigorously quantify and thus en-risk patient and public preferences. Using such instruments can reduce noise when studying public sentiment while simultaneously developing quantitative key performance indicators (KPIs) to track enthusiasm for product cycles, compare nascent and existing technologies, enhance narrative storytelling, and ultimately permit product de-risking. This can be done in a way that targets and modulates development and messaging according to discrete populations—in our case, on the basis of scaled vaccine or needle hesitancy (Aryal et al, 2024; Bridges et al, 2023).

Beyond the putative product development benefits, developing validated scales represents a prosocial contribution that yields outsized positive externalities, whether for academic institutions, start-ups, or established companies. Indeed, there is a virtuous feedback loop where instruments developed to accelerate internal product

iteration can be tapped as evidence of viability for funding agencies and even repurposed by other academic and commercial users, allowing for stronger funding applications while attracting partners, visibility, and financial support (Bridges et al, 2023). This may also help prime potential recipient populations to technologies and concepts, thereby lowering the familiarity barrier.

7.4 Summary: The ultrasound-mediated vaccine delivery model

Nowhere is the urgent need for better understanding of end user and public sentiment clearer than in vaccine development, where the combination of an apparently healthy population coupled with the potential for rare but high-profile serious side effects—in children, no less—can easily nudge a reasoned probabilistic risk-benefit trade-off into a technology that is perceived as dangerous.

In developing a transcutaneous ultrasound drug-device combination, making the narrative case required quantification of several key areas: needle fear, vaccine hesitancy, and preference for a nascent technology over needle and syringe. While some scales for the first two existed, none met a priori thresholds for multidimensionality and psychometric robustness. Three new instruments were therefore developed and validated on demographically representative populations with more than 3,000 participants in the UK and US: the Oxford Needle Experience (ONE) scale (Kantor et al, 2023), the Oxford Vaccine Hesitancy Scale (OVHS) (Kantor et al, 2024b), and the Oxford Benchmark Scale for Rating Vaccine Technologies (OBSRVT) (Kantor et al, 2025).

This approach produced a reusable toolkit for the rapid, iterative, and quantitative comparison of any vaccine-delivery technology to standard needle and syringe benchmarks. By developing broad scales, investigators can now rely on openly available validated scales compare to outcomes (using the OBSRVT) across a range of sub-populations with risk factors for adoption enthusiasm (using the ONE scale and OVHS). Subscales provide even greater granular detail to explore what aspects of a product are most associated with potential adoption, allowing for rapid prototype iteration while making a stronger narrative argument to funding agencies and the public.

7.5 From the MVP to the MDP

Just as basic science and clinical data drive product development forward in an open creative fashion (Yanai & Lercher, 2025), a parallel system of early-stage human factors research can be pursued to both increase funding opportunities and encourage adoption and venture success (Figure 7.2). Broadening the definition of target validation beyond the biological targets of drugs and devices to include stakeholder readiness can aid in developing the MDP (Paul et al, 2010).

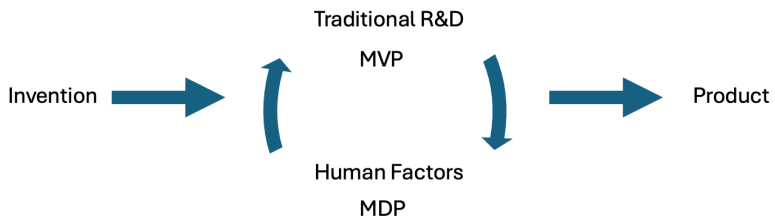


Figure 7.2 Traditional research and development (R&D) focuses on the minimum viable product (MVP), while incorporating human factors research early allows for a focus on the minimum desirable product (MDP).

Quantification may help mitigate the McNamara fallacy, where unquantified uncertainty is often ignored (Singh & Shah, 2023), though it may concomitantly introduce the risk of triggering Goodhart's law—once the measure becomes a target it ceases to be a reliable measure (Mattson et al, 2021). Still, by following a few key steps—en-risking and de-risking through scale development, early iterative scale use, and using quantitative data as KPIs for product development and narrative storytelling—research groups and firms can ultimately pivot from an MVP to an MDP, potentially speeding development, encouraging adoption, and enhancing the potential for venture success. Indeed, there is increasing recognition that quantitative human factors research can inform the development and regulatory evaluation of medical technologies: the US FDA's Patient Preference Initiative, launched in 2013, has incorporated patient preference information, including data from discrete choice experiments and benefit-risk trade-off surveys, into regulatory decisions for medical devices (Ho et al, 2015). The six-year IMI PREFER programme in Europe similarly developed frameworks for integrating patient preference studies into pharmaceutical decision-making across the product lifecycle (Janssens et al, 2023). These approaches highlight the value of systematic quantitative assessment of public and patient preferences, but they have largely been applied at the regulatory or post-development stage. Our approach—deploying validated instruments early in the development cycle to quantify otherwise unmeasured constructs such as needle fear, vaccine hesitancy, and technology preferences—seeks to shift this work upstream, enabling what we term "en-risking": the conversion of unmeasurable uncertainty into quantifiable, and therefore manageable, risk.

7.6 Future directions

Ultimately, this work can be used both to bolster the scientific and public buy-in rationale for pursuing ultrasound-mediated vaccination, while also serving more broadly as a model for future studies assessing public attitudes to vaccination, other novel biotechnology products, and beyond. In the near term, developing ultrasound-specific outcome measures may be helpful to gain further granular insight into patient and public motivations for preferring these technologies; such approaches can be particularly useful when iterating device prototypes or delivery approaches.

While this thesis includes an array of endpoints, investigating outcomes *not* related to self-reporting may be useful. Triangulating these responses and stated preference with revealed preferences and actual choices may further improve our ability to understand the nature and degree of preferences for given approaches while strengthening the validity argument for these instruments. Moreover, evaluating appropriate cut-offs based on ground truth outcomes of interest will provide additional meaningful data and further clarify appropriate approaches to using these instruments. Finally, further studies that involve invariance testing and investigating differential subgroup-level responses may be worthwhile in the future as well.

7.7 Conclusions

This thesis explored the ways in which instruments can be developed and validated to aid in developing alternative vaccine delivery technologies. After exploring the landscape of preexisting instruments, this thesis outlined the

development and validation of three scales that can be broadly applicable to developing these technologies: the Oxford Needle Experience (ONE) scale, providing an understanding of respondents attitudes to needles; the Oxford Vaccine Hesitancy Scale (OVHS), a multidimensional scale assessing vaccine hesitancy; and the Oxford Benchmark Scale for Rating Vaccine Technologies (OBSRVT), a scale that explicitly compares next-generation delivery technologies to a needle and syringe benchmark across a range of domains. Together, these scales represent a suite of instruments that can be used by researchers and clinicians beyond vaccine delivery technology research, highlighting the positive externalities associated with scale development. Using these tools iteratively may help research teams better understand preferences and de-risk technologies at an earlier stage, highlighting the value of the nascent technology to stakeholders while also permitting an early pivot if needed. In the short term, this work will hopefully permit the development of next-generation vaccine delivery platforms that are not only technically viable, but also acceptable, usable, and desirable by end users and the public.

References

Abd Halim, H., Abdul-Razak, S., Md Yasin, M. & Isa, M. R. (2020) Validation study of the Parent Attitudes About Childhood Vaccines (PACV) questionnaire: the Malay version. *Hum Vaccin Immunother*, 16(5), 1040-1049.

Abou-El-Enein, M., Duda, G. N., Gruskin, E. A. & Grainger, D. W. (2017) Strategies for Derisking Translational Processes for Biomedical Technologies. *Trends in Biotechnology*, 35(2), 100-108.

Alberts, J., Löwe, B., Glahn, M. A., Petrie, K., Laferton, J., Nestoriuc, Y. & Shedden-Mora, M. (2020) Development of the generic, multidimensional Treatment Expectation Questionnaire (TEX-Q) through systematic literature review, expert surveys and qualitative interviews. *BMJ Open*, 10(8), e036169.

Alver, C. G., Drabbe, E., Ishahak, M. & Agarwal, A. (2024) Roadblocks confronting widespread dissemination and deployment of Organs on Chips. *Nature Communications*, 15(1), 5118.

Anand, A., Molodecky, N. A., Pallansch, M. A. & Sutter, R. W. (2017) Immunogenicity to poliovirus type 2 following two doses of fractional intradermal inactivated poliovirus vaccine: A novel dose sparing immunization schedule. *Vaccine*, 35(22), 2993-2998.

Andersen, C. A., Brodersen, J., Rudbæk, T. R. & Jensen, M. B. (2021) Patients' experiences of the use of point-of-care ultrasound in general practice – a cross-sectional study. *BMC Family Practice*, 22(1), 116.

References

Armitage, E. P., Camara, J., Bah, S., Forster, A. S., Clarke, E., Kampmann, B. & de Silva, T. I. (2018) Acceptability of intranasal live attenuated influenza vaccine, influenza knowledge and vaccine intent in The Gambia. *Vaccine*, 36(13), 1772-1780.

Arnou, R., Eavis, P., De Juanes Pardo, J.-R., Ambrozaitis, A., Kazek, M.-P. & Weber, F. (2010) Immunogenicity, large scale safety and lot consistency of an intradermal influenza vaccine in adults aged 18-60 years Randomized, controlled, phase III trial. *Human Vaccines*, 6(4), 346-354.

Aryal, S., Blankenship, J. M., Bachman, S. L., Hwang, S., Zhai, Y., Richards, J. C., Clay, I. & Lyden, K. (2024) Patient-centricity in digital measure development: co-evolution of best practice and regulatory guidance. *npj Digital Medicine*, 7(1), 128.

Attwell, K., Smith, D. T. & Ward, P. R. (2018) 'The Unhealthy Other': How vaccine rejecting parents construct the vaccinating mainstream. *Vaccine*, 36(12), 1621-1626.

Bae, S. & Kim, H. (2023) [The Reliability and Validity of the Korean Version of the 5C Psychological Antecedents of Vaccination Scale]. *J Korean Acad Nurs*, 53(3), 324-339.

Baker, B., Hacker, E., Siller, G., Lee, M., Mursaliyev, N. & Forster, A. (2023) Evaluation of the self-administration potential of high-density microarray patches to human skin: A preliminary study. *Human Vaccines & Immunotherapeutics*, 19(1), 2189409.

Bani, M., Pepe, A., Zorzi, F., Ardenghi, S., Rampoldi, G. & Strepparava, M. G. (2020) Validation of the Italian version of the Injection Phobia Scale - Anxiety and application to blood donors. *Mediterranean Journal of Clinical Psychology*, 8(3).

Bardosh, K., de Figueiredo, A., Gur-Arie, R., Jamrozik, E., Doidge, J., Lemmens, T., Keshavjee, S., Graham, J. E. & Baral, S. (2022) The unintended consequences of COVID-

References

19 vaccine policy: why mandates, passports and restrictions may cause more harm than good. *BMJ Glob Health*, 7(5), e008684.

Berge, K. G., Agdal, M. L., Vika, M. & Skeie, M. S. (2016) High fear of intra-oral injections: prevalence and relationship to dental fear and dental avoidance among 10- to 16-yr-old children. *European Journal Of Oral Sciences*, 124(6), pp572-579.

Berger, M. N., Davies, C., Mathieu, E., Shaban, R. Z., Bag, S. & Skinner, S. R. (2024) Developing and validating a scale to measure the perceptions of safety, usability and acceptability of microarray patches for vaccination: a study protocol. *Therapeutic Advances in Vaccines and Immunotherapy*, 12.

Betsch, C., Schmid, P., Heinemeier, D., Korn, L., Holtmann, C. & Böhm, R. (2018) Beyond confidence: Development of a measure assessing the 5C psychological antecedents of vaccination. *PLOS One*, 13(12), e0208601.

Bienvenu, O. J. & Eaton, W. W. (1998) The epidemiology of blood-injection-injury phobia. *Psychological Medicine*, 28(5), 1129-1136.

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R. & Young, S. L. (2018) Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. *Front Public Health*, 6, 149.

Boivin, A., Lehoux, P., Lacombe, R., Burgers, J. & Grol, R. P. (2014) Involving patients in setting priorities for healthcare improvement: a cluster randomized trial. *Implement Sci*, 9(1), 24-24.

Boots Pharmacy (2025) *COVID-19 Vaccination Service*, 2025. Available at: https://www.boots.com/online/pharmacy-services/covid-19-vaccination-service [Accessed 02/12/2025].

References

Boragno, P., Fiabane, E., Colledani, D., Dalla Gasperina, D., Setti, I., Sommovigo, V. & Gabanelli, P. (2023) Attitude towards Intranasal Vaccines and Psychological Determinants: Effects on the General Population in Northern Italy. *Vaccines (Basel)*, 11(1), 138.

Brewer, N. T., Chapman, G. B., Rothman, A. J., Leask, J. & Kempe, A. (2017) Increasing Vaccination: Putting Psychological Science Into Action. *Psychological Science in the Public Interest*, 18(3), 149-207.

Bridges, J. F. P., de Bekker-Grob, E. W., Hauber, B., Heidenreich, S., Janssen, E., Bast, A., Hanmer, J., Danyliv, A., Low, E., Bouvy, J. C. & Marshall, D. A. (2023) A Roadmap for Increasing the Usefulness and Impact of Patient-Preference Studies in Decision Making in Health: A Good Practices Report of an ISPOR Task Force. *Value Health*, 26(2), 153-162.

Brown, K. F., Shanley, R., Cowley, N. A. L., van Wijgerden, J., Toff, P., Falconer, M., Ramsay, M., Hudson, M. J., Green, J., Vincent, C. A., Kroll, J. S., Fraser, G. & Sevdalis, N. (2011) Attitudinal and demographic predictors of measles, mumps and rubella (MMR) vaccine acceptance: Development and validation of an evidence-based measurement instrument. *Vaccine*, 29(8), 1700-1709.

Browne, M. W. & Cudeck, R. (1992) Alternative Ways of Assessing Model Fit. *Sociological Methods & Research*, 21(2), 230-258.

Bussink-Voorend, D., Hautvast, J. L. A., Wiersma, T., Akkermans, R. & Hulscher, M. (2025) Developing a practical tool for measuring parental vaccine hesitancy: A people-centered validation approach in Dutch. *Hum Vaccin Immunother*, 21(1), 2466303.

References

Bussink-Voorend, D. M., Hautvast, J. L., Vandeberg, L., Visser, O. & Hulscher, M. E. J. L. (2022) A systematic literature review to clarify the concept of vaccine hesitancy. *Nat Hum Behav*, 6(12), 1634-1648.

Cahn, D., Amosu, M., Maisel, K. & Duncan, G. A. (2023) Biomaterials for intranasal and inhaled vaccine delivery. *Nat Rev Bioeng*, 1(2), 83-84.

Callaghan, T., Motta, M., Sylvester, S., Trujillo, K. L. & Blackburn, C. C. (2019) Parent psychology and the decision to delay childhood vaccination. *Social Science & Medicine*, 238.

Callaghan, T., Washburn, D., Goidel, K., Nuzhath, T., Spiegelman, A., Scobee, J., Moghtaderi, A. & Motta, M. (2022) Imperfect messengers? An analysis of vaccine confidence among primary care physicians. *Vaccine*, 40(18), 2588-2603.

Caplan, A. L., Ferguson, K. & Williamson, A. (2024) Ethical Challenges of Advances in Vaccine Delivery Technologies. *The Hastings Center Report*, 54(1), 13-15.

Cappelleri, J. C., Lundy, J. J. & Hays, R. D. (2014) Overview of Classical Test Theory and Item Response Theory for the Quantitative Assessment of Items in Developing Patient-Reported Outcomes Measures Comment. *Clinical Therapeutics*, 36(5), 648-662.

Carpiano, R. M., Callaghan, T., DiResta, R., Brewer, N. T., Clinton, C., Galvani, A. P., Lakshmanan, R., Parmet, W. E., Omer, S. B., Buttenheim, A. M., Benjamin, R. M., Caplan, A., Elharake, J. A., Flowers, L. C., Maldonado, Y. A., Mello, M. M., Opel, D. J., Salmon, D. A., Schwartz, J. L., Sharfstein, J. M. & Hotez, P. J. (2023) Confronting the evolution and expansion of anti-vaccine activism in the USA in the COVID-19 era. *Lancet*, 401(10380), 967-970.

References

Carrion, M. L. (2017) “You need to do your research”: Vaccines, contestable science, and maternal epistemology. *Public Understanding of Science*, 27(3), 310-324.

Chadwick, A., Kaiser, J., Vaccari, C., Freeman, D., Lambe, S., Loe, B., Vanderslott, S., Lewandowsky, S., Conroy, M., Ross, A., Innocenti, S., Pollard, A., Waite, F., Larkin, M., Rosebrock, L., Jenner, L., McShane, H., Giubilini, A., Petit, A. & Yu, L. (2021) Online social endorsement and Covid-19 vaccine hesitancy in the United Kingdom. *Social Media + Society*, 7(2).

Cooper, S. & Wiysonge, C. S. (2023) Towards a More Critical Public Health Understanding of Vaccine Hesitancy: Key Insights from a Decade of Research. *Vaccines (Basel)*, 11(7), 1155.

Costello, A. B. & Osborne, J. (2005) Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, 10, 7.

de Figueiredo, A. & Larson, H. J. (2021) Exploratory study of the global intent to accept COVID-19 vaccinations. *Commun Med*, 1(1), 30.

de Figueiredo, A., Larson, H. J. & Reicher, S. D. (2021) The potential impact of vaccine passports on inclination to accept COVID-19 vaccinations in the United Kingdom: Evidence from a large cross-sectional survey and modeling study. *EClinicalMedicine*, 40, 101109-101109.

de Figueiredo, A., Simas, C., Karafillakis, E., Paterson, P. & Larson, H. J. (2020) Mapping global trends in vaccine confidence and investigating barriers to vaccine uptake: a large-scale retrospective temporal modelling study. *Lancet*, 396(10255), 898-908.

References

de Vet, H. C. W., Mokkink, L. B., Mosmuller, D. G. & Terwee, C. B. (2017) Spearman-Brown prophecy formula and Cronbach's alpha: different faces of reliability and opportunities for new applications. *J Clin Epidemiol*, 85, 45-49.

Deacon, B. & Abramowitz, J. (2006) Fear of needles and vasovagal reactions among phlebotomy patients. *Journal of Anxiety Disorders*, 20(7), 946-960.

DeWees, T. A., Mazza, G. L., Golafshar, M. A. & Dueck, A. C. (2020) Investigation Into the Effects of Using Normal Distribution Theory Methodology for Likert Scale Patient-Reported Outcome Data From Varying Underlying Distributions Including Floor/Ceiling Effects. *Value Health*, 23(5), 625-631.

Dinno, A. (2009) Implementing Horn's parallel analysis for principal component analysis and factor analysis. *Stata Journal*, 9(2), 291-298.

Dubé, È. & MacDonald, N. E. (2016) Managing the risks of vaccine hesitancy and refusals. *Lancet Infect Dis*, 16(5), 518-519.

Dubé, E., Vivion, M. & MacDonald, N. E. (2014) Vaccine hesitancy, vaccine refusal and the anti-vaccine movement: influence, impact and implications. *Expert review of Vaccines*, 14(1), 99-117.

Duchsherer, A., Jason, M., Platt, C. A. & Majdik, Z. P. (2020) Immunized against science: Narrative community building among vaccine refusing/hesitant parents. *Public Understanding of Science*, 29(4), 419-435.

Duplaga, M., Zwierczyk, U. & Kowalska-Duplaga, K. (2022) The Assessment of the Reliability and Validity of the Polish Version of the Adult Vaccine Hesitancy Scale (PL-aVHS) in the Context of Attitudes toward COVID-19 Vaccination. *Vaccines*, 10(10), 1666.

References

Durbach, N. (2000) 'They Might As Well Brand Us': Working-Class Resistance to Compulsory Vaccination in Victorian England. *Soc Hist Med*, 13(1), 45-63.

Eagan, R. L., Larson, H. J. & de Figueiredo, A. (2023) Recent trends in vaccine coverage and confidence: A cause for concern. *Hum Vaccin Immunother*, 19(2), 2237374.

Eisenblaetter, M., Madiouni, C., Laraki, Y., Capdevielle, D. & Raffard, S. (2023) Adaptation and Validation of a French Version of the Vaccination Attitudes Examination (VAX) Scale. *Vaccines (Basel)*, 11(5).

Eisinga, R., Grotenhuis, M. t. & Pelzer, B. (2013) The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? *International Journal of Public Health*, 58(4), 637-642.

Eskola, J., Duclos, P., Schuster, M., MacDonald, N. E. & the, S. W. G. o. V. H. (2015) How to deal with vaccine hesitancy? *Vaccine*, 33(34), 4215-4217.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C. & Strahan, E. J. (1999) Evaluating the Use of Exploratory Factor Analysis in Psychological Research. *Psychological Methods*, 4(3), 272-299.

Farag, N. H., Mansour, Z., Torossian, L., Said, R., Snider, C. J. & Ehrhardt, D. (2018) Feasibility of jet injector use during inactivated poliovirus vaccine house-to-house vaccination campaigns. *Vaccine*, 36(32), 4935-4938.

Fenta, E. T., Tiruneh, M. G., Delie, A. M., Kidie, A. A., Ayal, B. G., Limenh, L. W., Astatkie, B. G., Workie, N. K., yigzaw, Z. A., Bogale, E. k. & Anagaw, T. F. (2023) Health literacy and COVID-19 vaccine acceptance worldwide: A systematic review. *SAGE Open Med*, 9(11), 20503121231197869.

References

Fernando, G. J. P., Hickling, J., Flores, C. M. J., Griffin, P., Anderson, C. D., Skinner, S. R., Davies, C., Witham, K., Pryor, M., Bodle, J., Rockman, S., Frazer, I. H. & Forster, A. H. (2018) Safety, tolerability, acceptability and immunogenicity of an influenza vaccine delivered to human skin by a novel high-density microprojection array patch (Nanopatch (TM)). *Vaccine*, 36(26), 3779-3788.

Flood, E. M., Ryan, K. J., Rousculp, M. D., Beusterien, K. M., Block, S. L., Hall, M. C. & Mahadevia, P. J. (2011) A survey of children's preferences for influenza vaccine attributes. *Vaccine*, 29(26), 4334-4340.

Flora, D. B. & Curran, P. J. (2004) An Empirical Evaluation of Alternative Methods of Estimation for Confirmatory Factor Analysis With Ordinal Data. *Psychol Methods*, 9(4), 466-491.

Forster, A. & Junger, M. (2022) Opportunities and challenges for commercializing microarray patches for vaccination from a MAP developer's perspective. *Human Vaccines & Immunotherapeutics*, 18(4), 2050123.

Fredrikson, M., Annas, P., Fischer, H. & Wik, G. (1996) Gender and age differences in the prevalence of specific fears and phobias. *Behaviour Research and Therapy*, 34(1), 33-39.

Freeman, D., Lambe, S., Yu, L.-M., Freeman, J., Chadwick, A., Vaccari, C., Waite, F., Rosebrock, L., Petit, A., Vanderslott, S., Lewandowsky, S., Larkin, M., Innocenti, S., McShane, H., Pollard, A. & Loe, B. S. (2021) Injection fears and COVID-19 vaccine hesitancy. *Psychological Medicine*.

Freeman, D., Loe, B. S., Chadwick, A., Vaccari, C., Waite, F., Rosebrock, L., Jenner, L., Petit, A., Lewandowsky, S., Vanderslott, S., Innocenti, S., Larkin, M., Giubilini, A., Yu, L.

References

M., McShane, H., Pollard, A. J. & Lambe, S. (2022) COVID-19 vaccine hesitancy in the UK: the Oxford coronavirus explanations, attitudes, and narratives survey (Oceans) II. *Psychological Medicine*, 52(14), 3127-3141.

Frew, P. M., Murden, R., Mehta, C. C., Chamberlain, A. T., Hinman, A. R., Nowak, G., Mendel, J., Aikin, A., Randall, L. A., Hargreaves, A. L., Omer, S. B., Orenstein, W. A. & Bednarczyk, R. A. (2019) Development of a US trust measure to assess and monitor parental confidence in the vaccine system. *Vaccine*, 37(2), 325-332.

Frew, P. M., Paine, M. B., Rouphael, N., Schamel, J., Chung, Y., Mulligan, M. J. & Prausnitz, M. R. (2020) Acceptability of an inactivated influenza vaccine delivered by microneedle patch: Results from a phase I clinical trial of safety, reactogenicity, and immunogenicity. *Vaccine*, 38(45), 7175-7181.

Gao, Y., Liu, S., Xu, H., Wang, Y., Xu, G., Hu, F., Zhang, J. & Cai, Y. (2024) The Mpox Vaccine Hesitancy Scale for Mpox: Links with Vaccination Intention among Men Who Have Sex with Men in Six Cities of China. *Vaccines*, 12(9), 1009.

Garrido, L. E., Abad, F. J. & Ponsoda, V. (2011) Performance of Velicer's Minimum Average Partial Factor Retention Method With Categorical Variables. *Educational and Psychological Measurement*, 71(3), 551-570.

Gauna, F., Raude, J., Khouri, C., Cracowski, J.-L. & Ward, J. K. (2025) Exploring the relationship between experience of vaccine adverse events and vaccine hesitancy: A scoping review. *Hum Vaccin Immunother*, 21(1), 2471225.

Ghazy, R. M., Abd ElHafeez, S., Shaaban, R., Elbarazi, I., Abdou, M. S., Ramadan, A. & Kheirallah, K. A. (2021) Determining the Cutoff Points of the 5C Scale for Assessment

of COVID-19 Vaccines Psychological Antecedents among the Arab Population: A Multinational Study. *J Prim Care Community Health*, 12, 21501327211018568.

Gilkey, M. B., Magnus, B. E., Reiter, P. L., McRee, A. L., Dempsey, A. F. & Brewer, N. T. (2014) The Vaccination Confidence Scale: A brief measure of parents' vaccination beliefs. *Vaccine*, 32(47), 6259-6265.

Gilkey, M. B. P., Reiter, P. L. P., Magnus, B. E. M. A., McRee, A.-L. D., Dempsey, A. F. M. D. P. M. P. H. & Brewer, N. T. P. (2016) Validation of the Vaccination Confidence Scale: A Brief Measure to Identify Parents at Risk for Refusing Adolescent Vaccines. *Acad Pediatr*, 16(1), 42-49.

Gouglas, D., Thanh Le, T., Henderson, K., Kaloudis, A., Danielsen, T., Hammersland, N. C., Robinson, J. M., Heaton, P. M. & Røttingen, J.-A. (2018) Estimating the cost of vaccine development against epidemic infectious diseases: a cost minimisation study. *The Lancet Global Health*, 6(12), e1386-e1396.

Greiner, M. (1996) Two-graph receiver operating characteristic (TG-ROC): update version supports optimisation of cut-off values that minimise overall misclassification costs. *Journal of Immunological Methods*, 191(1), 93-94.

Griffin, P., Elliott, S., Krauer, K., Davies, C., Rachel Skinner, S., Anderson, C. D. & Forster, A. (2017) Safety, acceptability and tolerability of uncoated and excipient-coated high density silicon micro-projection array patches in human subjects. *Vaccine*, 35(48), 6676-6684.

Hacker, E., Baker, B., Lake, T., Ross, C., Cox, M., Davies, C., Skinner, S. R., Booy, R. & Forster, A. (2023) Vaccine microarray patch self-administration: An innovative

approach to improve pandemic and routine vaccination rates. *Vaccine*, 41(41), 5925-5930.

Hagedorn, B., Frey, K., Scarna, T. & El Sheikh, F. (2025) Microarray patches likely to reduce the operational costs of immunization: A Monte Carlo simulation study. *Vaccine*, 50, 126840-126840.

Hako, S., Kambara, K. & Ogata, A. (2022) The development and validation of the multidimensional fear-of-injection scale. *Health Psychology and Behavioral Medicine*, 10(1), 806-817.

Hanley, J. A. & McNeil, B. J. (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143(1), 29-36.

Hendrickson, A. E. & White, P. O. (1964) Promax: A quick method for rotation to oblique simple structure. *British Journal of Statistical Psychology*, 17(1), 65-70.

Hettinga, J. & Carlisle, R. (2020) Vaccination into the Dermal Compartment: Techniques, Challenges, and Prospects. *Vaccines (Basel)*, 8(3), 534.

Hettinga, J. K., Lyons, B., Balkaran, J., Rijal, P., Dunn-Lawless, D., Caproni, L., Gray, M., Suslick, K. S., Coussios, C. C. & Carlisle, R. C. (2023) Cavitation-Mediated Transcutaneous Delivery of Protein and Nucleotide-based Antigen for Rapid High-level Immune Responses. *Advanced Therapeutics*, 6(12), 2300102.

Hilário, A. P., Scavarda, A., Numerato, D., Mendonça, J., Cardano, M., Marhankova, J., Gariglio, L., Vuolanto, P., Anderson, A., Auvinen, P., Bracke, P., Douglass, T., Hobson-West, P., Lermytte, E., Polak, P. & Rudek, T. (2023) Recruiting a Hard-to-Reach, Hidden and Vulnerable Population: The Methodological and Practical Pitfalls of Researching Vaccine-Hesitant Parents. *Qualitative Health Research*, 33(13), 1189-1202.

References

Ho, M. P., Gonzalez, J. M., Lerner, H. P., Neuland, C. Y., Whang, J. M., McMurry-Heath, M., Brett Hauber, A. & Irony, T. (2015) Incorporating patient-preference evidence into regulatory decision making. *Surg Endosc*, 29(10), 2984-2993.

Holgado–Tello, F. P., Chacón–Moscoso, S., Barbero–García, I. & Vila–Abad, E. (2010) Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. *Quality & Quantity*, 44(1), 153-166.

Horn, J. L. (1965) A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179-185.

Howard, M. C. (2022) A more comprehensive measure of vaccine hesitancy: Creation of the Multidimensional Vaccine Hesitancy Scale (MVHS). *Journal of Health Psychology*, 27(10), 2402-2419.

Hu, L.-t. & Bentler, P. M. (1999) Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55.

Jackson, D. L., Gillaspy, J. A. & Purc-Stephenson, R. (2009) Reporting Practices in Confirmatory Factor Analysis: An Overview and Some Recommendations. *Psychol Methods*, 14(1), 6-23.

Jaiswal, J. & Halkitis, P. N. (2019) Towards a More Inclusive and Dynamic Understanding of Medical Mistrust Informed by Science. *Behavioral Medicine*, 45(2), 79-85.

Janssens, R., Barbier, L., Muller, M., Cleemput, I., Stoeckert, I., Whichello, C., Levitan, B., Hammad, T. A., Girvalaki, C., Ventura, J.-J., Bywall, K. S., Pinto, C. A., Schoefs, E., Katz, E. G., Kihlbom, U. & Huys, I. (2023) How can patient preferences be used and

communicated in the regulatory evaluation of medicinal products? Findings and recommendations from IMI PREFER and call to action. *Frontiers in Pharmacology*, 14, 1192770.

Jiang, X., Su, M.-H., Hwang, J., Lian, R., Brauer, M., Kim, S. & Shah, D. (2021) Polarization Over Vaccination: Ideological Differences in Twitter Expression About COVID-19 Vaccine Favorability and Specific Hesitancy Concerns. *Social Media + Society*, 7(3), 20563051211048413.

Johnson, N. F., Velásquez, N., Restrepo, N. J., Leahy, R., Gabriel, N., El Oud, S., Zheng, M., Manrique, P., Wuchty, S. & Lupu, Y. (2020) The online competition between pro- and anti-vaccination views. *Nature*, 582(7811), 230-233.

Kaaijk, P., Kleijne, D. E., Knol, M. J., Harmsen, I. A., Ophorst, O. J. A. E. & Rots, N. Y. (2014) Parents' attitude toward multiple vaccinations at a single visit with alternative delivery methods. *Human Vaccines & Immunotherapeutics*, 10(8), 2483-2489.

Kaiser, H. F. (1974) An Index of Factorial Simplicity. *Psychometrika*, 39(1), 31-36.

Kanekiyo, M., Ellis, D. & King, N. P. (2019) New Vaccine Design and Delivery Technologies. *The Journal of Infectious Diseases*, 219(Supplement_1), S88-S96.

Kantor, B. N. & Kantor, J. (2021) Development and validation of the Oxford Pandemic Attitude Scale-COVID-19 (OPAS-C): an internet-based cross-sectional study in the UK and USA. *BMJ Open*, 11(7), e043758-e043758.

Kantor, J. (2020) Behavioral epidemiology: vaccine hesitancy, sunscreen hesitancy, and smoking cessation. *Journal of the American Academy of Dermatology International*, 1(2), 222-223.

References

Kantor, J., Aasi, S. Z., Alam, M., Paoli, J. & Ratner, D. (2024a) Development and Validation of the Oxford Skin Cancer Treatment Scale, a Patient-Reported Outcome Measure for Health-Related Quality of Life and Treatment Satisfaction After Skin Cancer Treatment. *Dermatol Surg*, 50(11), 991-996.

Kantor, J., Carlisle, R. C., Morrison, M., Pollard, A. J. & Vanderslott, S. (2024b) Oxford Vaccine Hesitancy Scale (OVHS): a UK-based and US-based online mixed-methods psychometric development and validation study of an instrument to assess vaccine hesitancy. *BMJ Open*, 14(10), e084669.

Kantor, J., Carlisle, R. C., Vanderslott, S., Pollard, A. J. & Morrison, M. (2025) Development and validation of the Oxford Benchmark Scale for Rating Vaccine Technologies (OBSRVT), a scale for assessing public attitudes to next-generation vaccine delivery technologies. *Hum Vaccin Immunother*, 21(1), 2469994.

Kantor, J., Vanderslott, S., Morrison, M., Pollard, A. J. & Carlisle, R. C. (2023) The Oxford Needle Experience (ONE) scale: a UK-based and US-based online mixed-methods psychometric development and validation study of an instrument to assess needle fear, attitudes and expectations in the general public. *BMJ open*, 13(12), e074466.

Karabulut, N., Gurcayir, D., Yaman Aktas, Y. & Abi, O. (2022) The Turkish version of the Oxford Covid-19 Vaccine Hesitancy Scale. *Central European Journal of Nursing and Midwifery*, 13(3), 691-698.

Kleinknecht, R., Kleinknecht, E. E., Sawchuk, C. N., Lee, T. C. & Lohr, J. (1999) The Medical Fear Survey: Psychometric properties. *the Behavior Therapist*, 22, 109-112, 119.

References

Kleinknecht, R. A., Thorndike, R. M. & Walls, M. M. (1996) Factorial dimensions and correlates of blood, injury, injection and related medical fears: Cross validation of the medical fear survey. *Behaviour Research and Therapy*, 34(4), 323-331.

Klorman, R., Weerts, T. C., Hastings, J. E., Melamed, B. G. & Lang, P. J. (1974) Psychometric description of some specific-fear questionnaires. *Behavior Therapy*, 5(3), 401-409.

Koo, T. K. & Li, M. Y. (2016) A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med*, 15(2), 155-63.

Kose, S. & Mandiracioglu, A. (2007) Fear of blood/injection in healthy and unhealthy adults admitted to a teaching hospital. *International Journal of Clinical Practice*, 61(3), 453-457.

Kotta, I., Kalcza-Janosi, K., Szabo, K. & Marschalko, E. E. (2022) Development and Validation of the Multidimensional COVID-19 Vaccine Hesitancy Scale. *Human Vaccines & Immunotherapeutics*, 18(1), 1-10.

Kreps, S., Prasad, S., Brownstein, J. S., Hswen, Y., Garibaldi, B. T., Zhang, B. & Kriner, D. L. (2020) Factors Associated With US Adults' Likelihood of Accepting COVID-19 Vaccination. *JAMA Netw Open*, 3(10), e2025594-e2025594.

Kumar, M., Dogra, R. & Mandal, U. K. (2022a) Nanomaterial-based delivery of vaccine through nasal route: Opportunities, challenges, advantages, and limitations. *Journal of Drug Delivery Science and Technology*, 74, 103533.

Kumar, P., Bird, C., Holland, D., Joshi, S. B. & Volkin, D. B. (2022b) Current and next-generation formulation strategies for inactivated polio vaccines to lower costs,

increase coverage, and facilitate polio eradication. *Human Vaccines & Immunotherapeutics*, 18(7), 2154100.

Lalani, H. S., Nagar, S., Sarpatwari, A., Barenie, R. E., Avorn, J., Rome, B. N. & Kesselheim, A. S. (2023) US public investment in development of mRNA covid-19 vaccines: retrospective cohort study. *BMJ*, 380, e073747.

Lancsar, E. & Louviere, J. (2008) Conducting Discrete Choice Experiments to Inform Healthcare Decision Making: A User's Guide. *Pharmacoeconomics*, 26(8), 661-677.

Larson, H. J., Clarke, R. M., Jarrett, C., Eckersberger, E., Levine, Z., Schulz, W. S. & Paterson, P. (2018) Measuring trust in vaccination: A systematic review. *Human Vaccines & Immunotherapeutics*, 14(7), 1599-1609.

Larson, H. J., Jarrett, C., Schulz, W. S., Chaudhuri, M., Zhou, Y. Q., Dube, E., Schuster, M., MacDonald, N. E., Wilson, R. & Hesitancy, S. W. G. V. (2015) Measuring vaccine hesitancy: The development of a survey tool. *Vaccine*, 33(34), 4165-4175.

Lazarus, J. V., Ratzan, S. C., Palayew, A., Gostin, L. O., Larson, H. J., Rabin, K., Kimball, S. & El-Mohandes, A. (2020) A global survey of potential acceptance of a COVID-19 vaccine. *Nature Medicine*, 1-4.

Ledesma, R. D., Valero-Mora, P. & Macbeth, G. (2015) The Scree Test and the Number of Factors: a Dynamic Graphics Approach. *The Spanish Journal Of Psychology*, 18.

Leong, C., Jin, L., Kim, D., Kim, J., Teo, Y. Y. & Ho, T.-H. (2022) Assessing the impact of novelty and conformity on hesitancy towards COVID-19 vaccines using mRNA technology. *Communications Medicine*, 2(1), 61.

References

Li, S., Gong, T., Chen, G., Liu, P., Lai, X., Rong, H., Ma, X., Hou, Z. & Fang, H. (2022) Parental preference for influenza vaccine for children in China: a discrete choice experiment. *BMJ Open*, 12(6), e055725.

Liljequist, D., Elfving, B. & Skavberg Roaldsen, K. (2019) Intraclass correlation - A discussion and demonstration of basic features. *PLoS One*, 14(7), e0219854.

Lorenzo-Seva, U. & Ferrando, P. J. (2021) MSA: The Forgotten Index for Identifying Inappropriate Items Before Computing Exploratory Item Factor Analysis. *Methodology*, 17(4), 296-306.

Lutz, C. S., Fink, R. V., Cloud, A. J., Stevenson, J., Kim, D. & Fiebelkorn, A. P. (2020) Factors associated with perceptions of influenza vaccine safety and effectiveness among adults, United States, 2017–2018. *Vaccine*, 38(6), 1393-1401.

Ma, Y., Li, J., Cao, Y., Li, W., Shi, R., Jia, B., Wang, H., Yan, L., Suo, L., Yang, W., Wu, J. & Feng, L. (2023a) Acceptability for the influenza virus vector COVID-19 vaccine for intranasal spray: A cross-sectional survey in Beijing, China. *Hum Vaccin Immunother*, 19(2), 2235963.

Ma, Y., Li, J., Cao, Y., Li, W., Shi, R., Jia, B., Wang, H., Yan, L., Suo, L., Yang, W., Wu, J. & Feng, L. (2023b) Acceptability for the influenza virus vector COVID-19 vaccine for intranasal spray: A cross-sectional survey in Beijing, China. *Human Vaccines & Immunotherapeutics*, 19(2), 2235963.

MacDonald, N. E. & Sage Working Group on Vaccine Hesitancy (2015) Vaccine hesitancy: Definition, scope and determinants. *Vaccine*, 33(34), 4161-4164.

References

Machida, M., Takamiya, T., Odagiri, Y., Fukushima, N., Kikuchi, H. & Inoue, S. (2023) Estimation of cutoff score for the 7C of vaccination readiness scale. *Vaccine X*, 15, 100394.

Marti, M., de Cola, M., MacDonald, N. E., Dumolard, L. & Duclos, P. (2017) Assessments of global drivers of vaccine hesitancy in 2014—Looking beyond safety concerns. *PloS one*, 12(3), e0172310-e0172310.

Martin, L. R. & Petrie, K. J. (2017a) Understanding the Dimensions of Anti-Vaccination Attitudes: the Vaccination Attitudes Examination (VAX) Scale. *Ann Behav Med*, 51(5), 652-660.

Martin, L. R. & Petrie, K. J. (2017b) Understanding the Dimensions of Anti-Vaccination Attitudes: the Vaccination Attitudes Examination (VAX) Scale. *Annals of Behavioral Medicine*, 51(5), 652-660.

Marx, V. (2024) Collaborate, disseminate, accelerate. *Nature Methods*, 21(3), 372-376.

Mascherini, M. & Nivakoski, S. (2022) Social media use and vaccine hesitancy in the European Union. *Vaccine*, 40(14), 2215-2225.

Mathieu, E., Ritchie, H., Ortiz-Ospina, E., Roser, M., Hasell, J., Appel, C., Giattino, C. & Rodés-Guirao, L. (2021) A global database of COVID-19 vaccinations. *Nature Human Behaviour*, 5(7), 947-953.

Mattson, C., Bushardt, R. L. & Artino, A. R., Jr. (2021) "When a Measure Becomes a Target, It Ceases to be a Good Measure". *J Grad Med Educ*, 13(1), 2-5.

McLenon, J. & Rogers, M. A. M. (2019) The fear of needles: A systematic review and meta-analysis. *Journal of Advanced Nursing*, 75(1), 30-42.

References

McMurtry, C. M., Pillai Riddell, R., Taddio, A., Racine, N., Asmundson, G. J. G., Noel, M., Chambers, C. T., Shah, V., HelpinKids & Adults, T. (2015) Far From "Just a Poke": Common Painful Needle Procedures and the Development of Needle Fear. *The Clinical Journal of Pain*, 31(10 Suppl), S3-11.

McNeil, B. J., Pauker, S. G., Sox, H. C. & Tversky, A. (1982) On the Elicitation of Preferences for Alternative Therapies. *N Engl J Med*, 306(21), 1259-1262.

Melot, L., Thyfault, E., Hester, K., Prausnitz, M. R. & Bednarczyk, R. A. (2025) Perceptions of non-needle-based vaccination devices in the state of Georgia. *Vaccine*, 55, 127038.

Menon, I., Bagwe, P., Gomes, K. B., Bajaj, L., Gala, R., Uddin, M. N., D'Souza, M. J. & Zughaier, S. M. (2021) Microneedles: A New Generation Vaccine Delivery System. *Micromachines (Basel)*, 12(4), 435.

Miller, M. A. & Pisani, E. (1999) The cost of unsafe injections. *Bull World Health Organ*, 77(10), 808-811.

Mokken, R. J. & Lewis, C. (1982) A nonparametric approach to the analysis of dichotomous item responses. *Applied Psychological Measurement*, 6(4), 417-430.

Morgado, F. F. R., Meireles, J. F. F., Neves, C. M., Amaral, A. C. S. & Ferreira, M. E. C. (2017) Scale development: ten main limitations and recommendations to improve future research practices. *Psicol Reflex Crit*, 30(1), 3-20.

Mutlu, M., Cayir, Y. & Kasali, K. (2023) Validity and reliability of the Turkish version of the Parent Attitudes About Childhood Vaccines (PACV) scale. *J Healthc Qual Res*, 38(1), 11-19.

References

Nguyen, B. T., Brown, A. L., Jones, F., Jones, L., Withers, M., Ciesielski, K. M., Franks, J. M. & Wang, C. (2021a) "I'm not going to be a guinea pig:" Medical mistrust as a barrier to male contraception for Black American men in Los Angeles, CA. *Contraception (Stoneham)*, 104(4), 361-366.

Nguyen, T. T., Oh, Y., Kim, Y., Shin, Y., Baek, S.-K. & Park, J.-H. (2021b) Progress in microneedle array patch (MAP) for vaccine delivery. *Human Vaccines & Immunotherapeutics*, 17(1), 316-327.

Norman, J. J., Arya, J. M., McClain, M. A., Frew, P. M., Meltzer, M. I. & Prausnitz, M. R. (2014) Microneedle patches: Usability and acceptability for self-vaccination against influenza. *Vaccine*, 32(16), 1856-1862.

O'Shea, J., Prausnitz, M. R. & Rouphael, N. (2021) Dissolvable Microneedle Patches to Enable Increased Access to Vaccines against SARS-CoV-2 and Future Pandemic Outbreaks. *Vaccines (Basel)*, 9(4), 320.

Oduwole, E. O., Pienaar, E. D., Mahomed, H. & Wiysonge, C. S. (2022) Overview of Tools and Measures Investigating Vaccine Hesitancy in a Ten Year Period: A Scoping Review. *Vaccines (Basel)*, 10(8), 1198.

Okayasu, H., Sein, C., Blanc, D. C., Gonzalez, A. R., Zehrung, D., Jarrahian, C., Macklin, G. & Sutter, R. W. (2017) Intradermal Administration of Fractional Doses of Inactivated Poliovirus Vaccine: A Dose-Sparing Option for Polio Immunization. *Journal of Infectious Diseases*, 216, S161-S167.

Olatunji, B. O., Cisler, J. M., Deacon, B. J., Connolly, K. & Lohr, J. M. (2007a) The Disgust Propensity and Sensitivity Scale-Revised: Psychometric properties and

specificity in relation to anxiety disorder symptoms. *Journal of Anxiety Disorders*, 21(7), 918-930.

Olatunji, B. O., Ebesutani, C., Sawchuk, C. N., McKay, D., Lohr, J. M. & Kleinknecht, R. A. (2012) Development and Initial Validation of the Medical Fear Survey–Short Version. *Assessment*, 19(3), 318-336.

Olatunji, B. O., Sawchuk, C. N., Moretz, M. W., David, B., Armstrong, T. & Ciesielski, B. G. (2010) Factor Structure and Psychometric Properties of the Injection Phobia Scale-Anxiety. *Psychological Assessment*, 22(1), 167-179.

Olatunji, B. O., Smits, J. A. J., Connolly, K., Willems, J. & Lohr, J. A. (2007b) Examination of the decline in fear and disgust during exposure to threat-relevant stimuli in blood-injection-injury phobia. *Journal of Anxiety Disorders*, 21(3), 445-455.

Opel, D. J., Mangione-Smith, R., Taylor, J. A., Korfiatis, C., Wiese, C., Catz, S. & Martin, D. P. (2011a) Development of a survey to identify vaccine-hesitant parents The parent attitudes about childhood vaccines survey. *Human Vaccines*, 7(4), 419-425.

Opel, D. J., Taylor, J. A., Mangione-Smith, R., Solomon, C., Zhao, C., Catz, S. & Martin, D. (2011b) Validity and reliability of a survey to identify vaccine-hesitant parents. *Vaccine*, 29(38), 6598-6605.

Opel, D. J., Taylor, J. A., Zhou, C., Catz, S., Myaing, M. & Mangione-Smith, R. (2013) The relationship between parent attitudes about childhood vaccines survey scores and future child immunization status: a validation study. *JAMA Pediatr*, 167(11), 1065-71.

Orenstein, W. A., Gellin, B. G., Beigi, R. H., Despres, S., Lynfield, R., Maldonado, Y., Mouton, C., Rawlins, W., Rothholz, M. C., Smith, N., Thompson, K., Torres, C.,

References

Viswanath, K., Hosbach, P. & Natl Vaccine Advisory, C. (2015) Assessing the State of Vaccine Confidence in the United States: Recommendations from the National Vaccine Advisory Committee. *Public Health Reports*, 130(6), 573-595.

Öst, L.-G., Hellström, K. & Kåver, A. (1992) One versus five sessions of exposure in the treatment of injection phobia. *Behavior Therapy*, 23(2), 263-281.

Ost, L. G. (1992) BLOOD AND INJECTION PHOBIA - BACKGROUND AND COGNITIVE, PHYSIOLOGICAL, AND BEHAVIORAL VARIABLES. *Journal of Abnormal Psychology*, 101(1), 68-74.

Ovanessian, M. M., Fairbrother, N., Vorstenbosch, V., McCabe, R. E., Rowa, K. & Antony, M. M. (2019) Psychometric Properties and Clinical Utility of the Specific Phobia Questionnaire in an Anxiety Disorders Sample. *Journal of Psychopathology and Behavioral Assessment*, 41(1), 36-52.

Page, A. C., Bennett, K. S., Carter, O., Smith, J. & Woodmore, K. (1997a) The blood-injection symptom scale (BISS): Assessing a structure of phobic symptoms elicited by blood and injections. *Behaviour Research and Therapy*, 35(5), 457-464.

Page, A. C., Bennett, K. S., Carter, O., Smith, J. & Woodmore, K. (1997b) The Blood-Injection Symptom Scale (BISS): Assessing a structure of phobic symptoms elicited by blood and injections. *Behav Res Ther*, 35(5), 457-464.

Patakfalvi, L. & Benohanian, A. (2014) Needle-free anaesthesia, a promising option for the needle-phobic patient. *Br J Dermatol*, 170(5), 1191-1192.

Paul, S. M., Mytelka, D. S., Dunwiddie, C. T., Persinger, C. C., Munos, B. H., Lindborg, S. R. & Schacht, A. L. (2010) How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nature Reviews Drug Discovery*, 9(3), 203-214.

References

Peer, E., Brandimarte, L., Samat, S. & Acquisti, A. (2017) Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of experimental social psychology*, 70, 153-163.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y. & Podsakoff, N. P. (2003) Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. *J Appl Psychol*, 88(5), 879-903.

Qu, S., Zhou, M., Campy, K. S. & He, W. (2024) Predictors of parental acceptance to live attenuated influenza vaccine for children. *Human Vaccines & Immunotherapeutics*, 20(1), 2356343.

Radisic, G., Duncanson, E., Le Leu, R., Collins, K. L., Burke, A. L. J., Turner, J. K., Chur-Hansen, A., Donnelly, F., Hill, K., McDonald, S., Macauley, L. & Jesudason, S. (2022) Improving management of needle distress during the journey to dialysis through psychological education and training-the INJECT study feasibility pilot protocol. *Pilot Feasibility Stud*, 8(1), 28-28.

Rathje, S., He, J. K., Roozenbeek, J., Van Bavel, J. J. & van der Linden, S. (2022) Social media behavior is associated with vaccine hesitancy. *PNAS Nexus*, 1(4), 207.

Rodgers, A. M., Cordeiro, A. S. & Donnelly, R. F. (2019) Technology update: dissolvable microneedle patches for vaccine delivery. *Medical Devices-Evidence and Research*, 12, 379-398.

Rowan, N. & Wulff, D. (2007) Using qualitative methods to inform scale development. *Qualitative Report*, 12(3), 450-466.

References

Salmon, D. A., Dudley, M. Z., Glanz, J. M. & Omer, S. B. (2015a) Vaccine Hesitancy: Causes, Consequences, and a Call to Action. *American Journal of Preventive Medicine*, 49(6 Suppl 4), S391-S398.

Salmon, D. A., Dudley, M. Z., Glanz, J. M. & Omer, S. B. (2015b) Vaccine hesitancy: Causes, consequences, and a call to action. *Vaccine*, 33 Suppl 4, D66-D71.

Sanford, K. & Clifton, M. (2022) The Medical Mistrust Multiformat Scale: Links with vaccine hesitancy, treatment adherence, and patient-physician relationships. *Psychol Assess*, 34(1), 10-20.

Sarathchandra, D., Navin, M. C., Largent, M. A. & McCright, A. M. (2018) A survey instrument for measuring vaccine acceptance. *Preventive Medicine*, 109, 1-7.

Sathanapally, H., Sidhu, M., Fahami, R., Gillies, C., Kadam, U., Davies, M. J., Khunti, K. & Seidu, S. (2020) Priorities of patients with multimorbidity and of clinicians regarding treatment and health outcomes: a systematic mixed studies review. *BMJ Open*, 10(2), e033445.

Satorra, A. & Bentler, P. M. (2010) Ensuring Positiveness of the Scaled Difference Chi-square Test Statistic. *Psychometrika*, 75(2), 243-248.

Sawchuk, C. N., Lohr, J. M., Tolin, D. F., Lee, T. C. & Kleinknecht, R. A. (2000) Disgust sensitivity and contamination fears in spider and blood-injection-injury phobias. *Behaviour Research and Therapy*, 38(8), 753-762.

Sawchuk, C. N., Meunier, S. A., Lohr, J. M. & Westendorf, D. H. (2002) Fear, disgust, and information processing in specific phobia: The application of signal detection theory. *Journal of Anxiety Disorders*, 16(5), 495-510.

References

Schnyder, J. L., Garcia Garrido, H. M., De Pijper, C. A., Daams, J. G., Stijnis, C., Goorhuis, A. & Grobusch, M. P. (2021) Comparison of equivalent fractional vaccine doses delivered by intradermal and intramuscular or subcutaneous routes: A systematic review. *Travel Med Infect Dis*, 41, 102007-102007.

Seah, B. C.-Q. & Teo, B. M. (2018) Recent advances in ultrasound-based transdermal drug delivery. *International Journal of Nanomedicine*, 13, 7749-7763.

Shapiro, G. K., Holding, A., Perez, S., Amsel, R. & Rosberger, Z. (2016) Validation of the vaccine conspiracy beliefs scale. *Papillomavirus Res*, 2, 167-172.

Shapiro, G. K., Kaufman, J., Brewer, N. T., Wiley, K., Menning, L. & Leask, J. (2021) A critical review of measures of childhood vaccine confidence. *Curr Opin Immunol*, 71, 34-45.

Shapiro, G. K., Tatar, O., Dube, E., Amsel, R., Knauper, B., Naz, A., Perez, S. & Rosberger, Z. (2018) The vaccine hesitancy scale: Psychometric properties and validation. *Vaccine*, 36(5), 660-667.

Shelley-Egan, C., Bowman, D. M. & Robinson, D. K. R. (2018) Devices of Responsibility: Over a Decade of Responsible Research and Innovation Initiatives for Nanotechnologies. *Sci Eng Ethics*, 24(6), 1719-1746.

Shrout, P. E. & Fleiss, J. L. (1979) Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*, 86(2), 420-8.

Singh, T. & Shah, N. (2023) Competency-based medical education and the McNamara fallacy: Assessing the important or making the assessed important? *J Postgrad Med*, 69(1), 35-40.

References

Smith, L. E., Webster, R. K., Weinman, J., Amlôt, R., Yiend, J. & Rubin, G. J. (2017) Psychological factors associated with uptake of the childhood influenza vaccine and perception of post-vaccination side-effects: A cross-sectional survey in England. *Vaccine*, 35(15), 1936-1945.

Stoeckel, F., Carter, C., Lyons, B. A. & Reifler, J. (2022) The politics of vaccine hesitancy in Europe. *European Journal of Public Health*, 32(4), 636-642.

Stout, W. F. (1990) A new item response theory modeling approach with applications to unidimensionality assessment and ability estimation. *Psychometrika*, 55(2), 293-325.

Streiner, D. L. (2003) Starting at the Beginning: An Introduction to Coefficient Alpha and Internal Consistency. *Journal of Personality Assessment*, 80(1), 99-103.

Taddio, A., McMurtry, C. M., Logeman, C., Gudzak, V., de Boer, A., Constantin, K., Lee, S., Moline, R., Uleryk, E., Chera, T., MacDonald, N. E. & Pham, B. (2022) Prevalence of pain and fear as barriers to vaccination in children – Systematic review and meta-analysis. *Vaccine*, 40(52), 7526-7537.

Tang, J., Cai, L., Xu, C., Sun, S., Liu, Y., Rosenecker, J. & Guan, S. (2022) Nanotechnologies in Delivery of DNA and mRNA Vaccines to the Nasal and Pulmonary Mucosa. *Nanomaterials (Basel, Switzerland)*, 12(2), 226.

Terwee, C. B., Bot, S. D. M., de Boer, M. R., van der Windt, D. A. W. M., Knol, D. L., Dekker, J., Bouter, L. M. & de Vet, H. C. W. (2007) Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, 60(1), 34-42.

References

Tezel, A., Paliwal, S., Shen, Z. & Mitragotri, S. (2005) Low-frequency ultrasound as a transcutaneous immunization adjuvant. *Vaccine*, 23(29), 3800-3807.

The Lancet Child & Adolescent Health (2019) Vaccine hesitancy: a generation at risk. *The Lancet Child & Adolescent Health*, 3(5), 281.

Tram, K. H., Saeed, S., Bradley, C., Fox, B., Eshun-Wilson, I., Mody, A. & Geng, E. (2022) Deliberation, Dissent, and Distrust: Understanding Distinct Drivers of Coronavirus Disease 2019 Vaccine Hesitancy in the United States. *Clinical Infectious Diseases*, 74(8), 1429-1441.

Tversky, A. & Kahneman, D. (1981) The Framing of Decisions and the Psychology of Choice. *Science*, 211(4481), 453-458.

Ünal, İ., Ünal, E., Sertdemir, Y. & Kobaner, M. (2025) Defining optimal cut-off points for multiple class ROC analysis: generalization of the Index of Union method. *J Biopharm Stat*, 1-19.

van Overveld, M., de Jong, P. J. & Peters, M. L. (2011) The Multi-Dimensional Blood/Injury Phobia Inventory: Its psychometric properties and relationship with disgust propensity and disgust sensitivity. *J Anxiety Disord*, 25(3), 319-325.

van Overveld, W. J. M., de Jong, P. J., Peters, M. L., Cavanagh, K. & Davey, G. C. L. (2006) Disgust propensity and disgust sensitivity: Separate constructs that are differentially related to specific fears. *Personality and Individual Differences*, 41(7), 1241-1252.

van Tulleken, C., Rollins, N. & Coombes, R. (2024) Conflicts of interest: moving towards zero tolerance. *BMJ*, 387, q2574.

References

Wallis, J., Shenton, D. P. & Carlisle, R. C. (2019) Novel approaches for the design, delivery and administration of vaccine technologies. *Clinical and Experimental Immunology*, 196(2), 189-204.

Wang, E., Baras, Y. & Buttenheim, A. M. (2015) "Everybody just wants to do what's best for their child": Understanding how pro-vaccine parents can support a culture of vaccine hesitancy. *Vaccine*, 33(48), 6703-6709.

Ward, J. K., Gauna, F., Gagneux-Brunon, A., Botelho-Nevers, E., Cracowski, J.-L., Khouri, C., Launay, O., Verger, P. & Peretti-Watel, P. (2022) The French health pass holds lessons for mandatory COVID-19 vaccination. *Nature Medicine*, 28(2), 232-235.

Ward, P. R., Attwell, K., Meyer, S. B., Rokkas, P. & Leask, J. (2017) Understanding the perceived logic of care by vaccine-hesitant and vaccine-refusing parents: A qualitative study in Australia. *PLOS ONE*, 12(10), e0185955.

Watkins, M. W. (2022) *A step-by-step guide to exploratory factor analysis with Stata*. New York, NY: Routledge.

Watson, D., Clark, L. A. & Tellegen, A. (1988) Development and validation of brief measures of positive and negative affect - the PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063-1070.

Wenzel, A. & Holt, C. S. (2003) Validation of the Multidimensional Blood/Injury Phobia Inventory: Evidence for a unitary construct. *Journal of Psychopathology and Behavioral Assessment*, 25(3), 203-211.

Wilhelm, M., Bender, F. L., Euteneuer, F., Salzmann, S., Ewen, A.-C. I. & Rief, W. (2024) Psychometric validation of the Vaccination Attitudes Examination (VAX) scale in German pre-pandemic and mid-pandemic samples. *Scientific Reports*, 14(1), 30543.

References

Williams, S. C. P. (2013) Under the skin of intradermal vaccines. *Proc Natl Acad Sci U S A*, 110(25), 10049-10051.

World Health Organization (2019) *Ten threats to global health in 2019*, 2019. Available at: https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019 [Accessed 10/25/2020].

World Health Organization (2022) Understanding the behavioural and social drivers of vaccine uptake WHO position paper – May 2022. *Weekly Epidemiological Record*, 97(20), 209-224.

Wu, H. & Leung, S.-O. (2017) Can Likert Scales be Treated as Interval Scales?—A Simulation Study. *Journal Of Social Service Research*, 43(4), pp527-532.

Yanai, I. & Lercher, M. J. (2025) Openness guides discovery. *Nat Biotechnol*, 43(5), 667-668.

Yao, X., Peng, J., Ma, Y., Fu, M., Fang, Y., Fang, H. & Jiang, M. (2023) Development of a vaccine hesitancy scale for childhood immunization in China. *Vaccine*, 41(46), 6870-6877.

Young, S. L., Boateng, G. O., Jamaluddine, Z., Miller, J. D., Frongillo, E. A., Neilands, T. B., Collins, S. M., Wutich, A., Jepson, W. E., Stoler, J. & Network, H. R. C. (2019) The Household Water InSecurity Experiences (HWISE) Scale: development and validation of a household water insecurity measure for low-income and middle-income countries. *BMJ Global Health*, 4(5).

Yuan, J., Li, L., Dong, M., So, H. C., Cowing, B. J., Ip, D. K. M. & Liao, Q. (2024) Parental vaccine hesitancy and influenza vaccine type preferences during and after the COVID-19 Pandemic. *Commun Med (Lond)*, 4(1), 165-10.

References

Zucoloto, M. L., Gonçalez, T., Menezes, N. P., McFarland, W., Custer, B. & Martinez, E. Z. (2019) Fear of blood, injections and fainting as barriers to blood donation in Brazil. *Vox Sanguinis*, 114(1), pp38-46.

Zweig, M. H. & Campbell, G. (1993) Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin Chem*, 39(4), 561-77.

Appendix: Glossary of Psychometric Terms

Term	Definition
Face validity	The degree to which a scale appears, on its surface (face), to actually measure what it is intended to measure.
Content validity	The degree to which the items in a scale adequately represent the full range of the construct being measured.
Convergent validity	Evidence that a scale correlates with other measures of (theoretically) related constructs.
Discriminant (divergent) validity	Evidence that a scale does *not* correlate strongly with measures of (theoretically) unrelated constructs.
Construct validity	The overarching degree to which a scale measures the theoretical construct it is intended to measure; this broad terms encompasses convergent, discriminant, and structural validity as outlined above.
Predictive validity	Evidence that a scale score predicts a relevant future (classically) or concurrent outcome.
Test-retest reliability	The temporal stability of scale scores when administered to the same respondents at two timepoints; this assumes no true change in the underlying construct or attitudes to the construct.