Exploring the Value of Shared Clinical Trial Data: Policy, Application, and Impact

By

Natansh Deepak Modi
B. Pharm (Hons), B. Pharm Sci

Thesis
Submitted to Flinders University
for the degree of

Doctor of Philosophy
College of Medicine and Public Health
5/11/2024

TABLE OF CONTENTS

ABSTRACT

Clinical trials are pivotal for the evolution of modern medicine, yet the full potential of clinical trial data remains unrealised due to industry-developed data sharing policies that restrict accessibility. This thesis advocates that the democratisation of data sharing practices - making clinical trial data broadly accessible to independent and qualified researchers - is not only important for advancing medical care but also fulfills a moral duty to honour the contributions of clinical trial participants.

Chapter 1 is the introduction that positions the thesis within a global context by discussing the role of data transparency in fostering trust and accountability in medical research. This chapter sets the stage for the thesis's exploration of transparency issues across the pharmaceutical industry.

Chapter 2 audits the individual patient data (IPD) availability for anticancer medications, revealing low IPD transparency rates, especially for high revenue generating oncology medicines. This chapter raises questions about ethical obligations versus commercial interests.

Chapter 3 evaluates the transparency rates for clinical study reports and IPD in oncology within oncology and proposes actionable, evidence-based strategies to enhance data sharing practices, advocating for industry-wide reform.

Chapter 4 expands the scope to examine IPD transparency for the top revenue-generating drugs across all therapeutic categories, analysing whether oncology's data sharing issues are symptomatic of broader, industry-wide trends, potentially affecting other therapeutic areas.

Chapter 5, examines the pharmaceutical industry's 2013 commitments to transparency, juxtaposing them against current practices to identify gaps and propose essential updates fostering a culture of open data.

Chapter 6 moves to practical applications, detailing the creation of a predictive tool for neutropenia, a side effect of the anticancer drug Abemaciclib, using shared trial data to provide clinicians and patients with insights into differing risks and aiding personalised treatment decisions.

Chapter 7 examines the prognostic value of patient-reported outcomes (PROs) in advanced breast cancer, arguing that integrating PROs with clinical data could offer a more comprehensive approach to patient care, enhancing quality of life assessments alongside traditional metrics.

Chapter 8 synthesises the findings, offering a discussion of key insights, future directions, and the thesis's implications for global data sharing practices. It emphasises the need for ongoing collaboration between industry, regulatory agencies, and researchers to ensure clinical trial data serves its intended purpose.

DECLARATION

I certify that this thesis:

1. does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university

2. and the research within will not be submitted for any other future degree or diploma without the permission of Flinders University; and

3. to the best of my knowledge and belief, does not contain any material previously published or written by another person except where due reference is made in the text.

Signed:

Date: 5/11/2024

ACKNOWLEDGEMENTS

First and foremost, I want to acknowledge my parents, Deepak and Vaibhavi Modi. I thank them for instilling the values of hard work, perseverance, and intellectual curiosity in me. Their belief in my abilities has been the driving force – especially during times of self-doubt. I hope I can continue to make you both proud. Hiya, my dearest sister - you have been my sounding board, voice of reason, and bringer of endless laughs. I value your company, deeply.

My incredible friends - you brought cheerfulness to my life during some stressful moments. All of you helped me in finding the balance between hard work and enjoying the simple joys. Our adventures together and shared laughs were the reprieve I needed from the chaos of my PhD.

I have been extremely fortunate to be supervised by Professors McKinnon, Rowland, and Sorich. Their knowledge, guidance and collaborative spirit cultivated an environment where I could question, explore, and grow as a young researcher.

Words cannot capture the depths of my gratitude for Associate Professor Ashley Hopkins, my primary supervisor and mentor. Your keen intellect is matched only by your exceptional ability to communicate expectations while remaining open to new ideas. You celebrated my successes as your own, whilst also pushing me to overcome setbacks. I will carry your lessons forward, always.

I am deeply grateful to the National Health and Medical Research Council for funding the postgraduate scholarship, which provided me with a living stipend and made it possible for me to pursue my PhD. I am also thankful to the Flinders Health and Medical Research Institute for providing me with a top-up scholarship.

To anyone else who supported me directly or indirectly throughout this - thank you. This work would not have been possible without you.

Thank you all for being part of this chapter of my life.

LIST OF FIGURES

LIST OF TABLES

HDR COMPLETION PLAN

Chapter	Title	Status
2	Audit of data sharing by pharmaceutical companies for anticancer medicines approved by the US Food and Drug Administration	Published
3	Clinical study report and individual participant data transparency for FDA-approved anticancer drugs: A call for systematic data availability	Published
4	The state of individual participant data sharing for highest revenue medicines	Published
5	A 10-year update to the principles for clinical trial data sharing by pharmaceutical companies: perspectives based on a decade of literature and policies	Published
6	Prediction of severe neutropenia and diarrhoea in breast cancer patients treated with abemaciclib	Published
7	Patient-reported outcomes predict survival and adverse events following anticancer treatment initiation in advanced HER2-positive breast cancer	Published

ACRONYMS AND ABBREVIATIONS

ABC: Advanced Breast Cancer
BMI: Body Mass Index
BMJ: British Medical Journal
BC: Breast Cancer
CDK: Cyclin-Dependent Kinase
CI: Confidence Intervals
CSDR: Clinical Study Data Request.com
CSR: Clinical Study Report
ECOG PS: Eastern Cooperative Oncology Group Performance Status
EFPIA: European Federation of Pharmaceutical Industries and Associations
EMA: European Medicines Agency
ESMO: European Society for Medical Oncology
EU: European Union
EU CTR: European Union Clinical Trials Regulation
FAIR: Findable, Accessible, Interoperable and Reusable
FACT-B: Functional Assessment of Cancer Therapy – Breast
FDA: Food and Drug Administration
HER2: Human Epidermal Growth Factor Receptor 2
HR: Hazard Ratios
HR+/HER2-: Hormone Receptor-Positive/Human Epidermal Growth Factor 2-Negative
ICMJE: International Committee of Medical Journal Editors
IPD: Individual Participant Data
NCI CTCAE: National Cancer Institute Common Terminology Criteria for Adverse Events
NCT: National Clinical Trial
NIH: National Institutes of Health
NSAI: Non-Steroidal Aromatase Inhibitor
OS: Overall Survival
PhRMA: Pharmaceutical Research and Manufacturers of America
PFS: Progression-Free Survival
PROs: Patient-Reported Outcomes
RECIST: Response Evaluation Criteria in Solid Tumours
RCT: Randomised Controlled Trial
SAP: Statistical Analysis Plan
T-DM1: Ado-Trastuzumab Emtansine
US: United States
WBC: White Blood Cell
WHO: World Health Organization
c-statistic: Concordance Statistic

CHAPTER 1: INTRODUCTION

Clinical trials are fundamental to the progression of modern medicine, serving as the cornerstone for the development and validation of new treatments and interventions [1]. These trials assess the safety, efficacy, and overall risk-benefit profile of medical interventions, providing the data necessary for regulatory bodies to approve them for patient use [1]. Beyond their primary purpose of evaluating new treatments, clinical trials generate vast amounts of data that holds significant potential for advancing medical knowledge and improving patient outcomes [2]. Data, in the context of clinical trials, encompasses various forms, including study protocols, summary results, Clinical Study Reports (CSRs), and, importantly, Individual Participant Data (IPD). IPD encompasses detailed information on patient demographics, laboratory measurements, adverse events, and responses to treatment, as recorded during each clinical [2]. Thus, IPD provides the scientific community with rich, granular insights that can drive further advancements in medicine [2]. For instance, reanalysing IPD from clinical trials allows researchers to explore new hypotheses that were not part of the original study design, perform meta-analyses to synthesise broader evidence, or even develop personalised treatment models based on patient subgroups [3-5]. Therefore, in addition to their role in regulatory approvals, clinical trial data sharing plays an important role in advancing transparency and collaboration through the dissemination and accessibility of IPD. The movement toward open data sharing reflects a growing recognition that shared data leads to greater accountability, reproducibility, and efficiency [6]. This thesis aims to advocate for a data sharing framework that transforms IPD from static records into actionable insights that contribute to better patient outcomes and informed policy decisions. Such framework would enable researchers, clinicians, and policymakers to leverage IPD for a variety of purposes: optimising patient care, developing predictive tools, and refining clinical guidelines based on real-world evidence.

The push toward data transparency in clinical trials gained momentum over the past two decades [7]. One of the most visible achievements in this space has been the establishment of clinical trial registries, such as ClinicalTrials.gov and the European Union (EU) Clinical Trials Register, which provide public access to key details about ongoing and completed clinical trials [8, 9]. These registries aim to reduce publication bias by ensuring that all trials - not just those with positive outcomes - are registered and their basic results disclosed [10]. Trial registration has become a standard expectation for clinical research, demonstrating that transparency initiatives can be successfully implemented with the right policies and incentives in place. However, while trial registration has become standard practice, IPD sharing remains inconsistent [11]. The COVID-19 pandemic has further highlighted the importance of data sharing in medicine. During the global health crisis, researchers, governments, and pharmaceutical companies collaborated on an unprecedented

scale to share information, rapidly developing vaccines and treatments that saved millions of lives. Data sharing initiatives, such as the release of interim clinical trial results for COVID-19 vaccines, allowed regulatory agencies and healthcare providers to make timely decisions about vaccine deployment.

Within the pharmaceutical space, data sharing policies are designed to enhance the accessibility of clinical trial data (in this context and for the thesis, 'data' refers specifically to IPD) [12]. Building on the success of trial registries, the pharmaceutical industry made a landmark commitment in 2013 to share anonymised IPD for approved medicines upon request by qualified researchers [13]. Many companies (such as Roche, Pfizer, Lilly) have since become contributors to various data sharing platforms (such as Vivli, CSDR) to facilitate independent researcher access to IPD [14-16]. However, navigating these platforms has been shown to be time-consuming and complex, with researchers facing lengthy approval processes and restrictive data-use agreements [14]. The variability in data sharing practices across companies presents another significant challenge. Some pharmaceutical companies have been proactive in making their data available, while others remain reluctant, citing proprietary concerns and the potential misuse of data [17]. Additionally, access to IPD is often limited to researchers who can demonstrate a specific, pre-approved research question, leaving little room for exploratory analyses or secondary investigations that might yield unexpected insights [3]. These barriers stifle innovation and limit the impact of shared data and thus prevents the scientific community from fully leveraging the information generated in clinical trials.

Another persistent issue is the lack of standardisation in data sharing policies and platforms [14, 18]. While some platforms provide access to raw IPD, others offer only aggregated data or require researchers to use specific software to analyse datasets within a controlled environment [3]. This fragmentation complicates efforts to conduct large-scale meta-analyses or combine datasets from multiple sources. Moreover, differences in national and regional data privacy regulations introduce further complexity. These regulations, while essential for protecting patient privacy, can create uncertainty about what data can be shared and under what conditions further adding to the operational burden on researchers and data providers [19].

While commitments to data sharing represent progress, the gap between policy and practice remains a significant hurdle. Policy audits in research play an important role in bridging the gap between intentions and practice [20]. Auditing offers a structured method of evaluating how effectively pharmaceutical companies are applying data sharing initiatives. It can identify discrepancies between the commitments made by the companies and the actual accessibility of data on platforms. Just as financial audits expose weaknesses in governance and enable corrective measures, data sharing audits will further highlight bottlenecks that undermine the utility of shared data [21]. These

audits provide not just accountability but also a feedback loop for continuous improvement which will ensure that data sharing practices evolve in response to emerging challenges. Moreover, the auditing framework can drive positive behavioural change among stakeholders by encouraging greater transparency and advancing collaboration. When pharmaceutical companies see that their efforts are measured and recognised through policy audits, they are more likely to engage proactively in data sharing initiatives. This thesis builds on these principles by proposing a framework for auditing data sharing practices of pharmaceutical companies, using oncology trials and high-revenue pharmaceutical medicines as case studies. Through a detailed evaluation of IPD accessibility, the thesis also aims to identify specific barriers and develop practical recommendations for improving data sharing practices. Therefore, the thesis positions itself as both an audit and advocacy effort. It aims to emphasise the importance of ongoing evaluation, stakeholder collaboration, and practical policy reform which will set the stage for meaningful improvements in data sharing practices. The work presented in here acknowledges that achieving these goals will require sustained effort and engagement from multiple stakeholders. However, it also argues that the benefits - improved patient outcomes, accelerated innovation, and increased public trust - make these efforts worthwhile.

The latter part of this thesis explores practical applications which is an important step in demonstrating the full value of data sharing initiatives. Utilising shared IPD to real-world scenarios completes the cycle of transparency - progressing from policy and accessibility discussions to measurable impacts on patient care and clinical decision-making. This approach validates the importance of accessible IPD, illustrating how transparency initiatives can lead to improved patient outcomes, evidence-based policies, and faster innovation.

This thesis, "Exploring the Value of Shared Clinical Trial Data: Policy, Application, and Impact," aims to evaluate and address the multifaceted aspects of data sharing. The decision to focus on oncology in the first two chapters of this thesis stems from the important role cancer research plays in the advancement of modern medicine. Cancer, as a global health priority, demands diligent investigation. The data generated from these trials are important for the development of new therapies, refining treatment protocols, and advancing personalised medicine. However, despite the importance of this data, access to it remains inconsistent [11]. Therefore, Chapter 2 investigates the accessibility of IPD from oncology trials that supported the Food and Drug Administration approvals of new anti-cancer medications over the past decade. Chapter 3 explores the current landscape and future prospects of Clinical Study Reports and IPD sharing, acknowledging the gaps in their use within systematic reviews and meta-analyses.

Expanding the scope beyond oncology, Chapter 4 investigates the availability of IPD from the top 30 pharmaceutical products by revenue in 2021. This analysis offers a broader perspective on data sharing practices across different therapeutic areas, highlighting whether high-revenue products - often guarded closely due to proprietary interests - are subject to the same transparency standards as oncology therapies. This comparison allows the thesis to identify patterns in data accessibility and provide recommendations tailored to different therapeutic fields. Chapter 5 serves as the centrepiece of this thesis. It presents a detailed review that also provides policy updates and recommendations. The placement of the perspective piece in Chapter 5, after the empirical investigations, is a deliberate structural decision that reflects the dynamic and evolving nature of data transparency itself. Rather than positioning the literature review at the front of the thesis, Chapter 5 draws from the real-world insights uncovered in Chapters 2 through 4 to critique the existing policies more effectively. This allows the literature review to move beyond simply summarising the state of the field. Thus, it offers a fully contextualised evaluation of the current landscape. This chapter also incorporates insights from a multidisciplinary team of experts, including policymakers, clinicians, and data scientists. These stakeholders offer diverse perspectives on the challenges and opportunities associated with data sharing, ensuring that the recommendations presented are grounded in practical realities. Their input is essential for understanding the subtle trade-offs between transparency, privacy, and proprietary interests, as well as for identifying strategies that balance these competing priorities.

Building upon the audit and advocacy efforts outlined, the latter chapters of this thesis transition to demonstrating the tangible benefits and practical applications of accessible IPD. Leveraging shared clinical trial data, these chapters illustrate how IPD can be transformed into actionable insights that directly impact patient care and advance medical knowledge. For instance, Chapter 6 details the development of a predictive tool for neutropenia using shared IPD, enabling clinicians and patients to make more informed decisions regarding the use of Abemaciclib. Chapter 7 explores the integration of patient-reported outcomes in advanced breast cancer treatment, highlighting how IPD enhances our understanding of patient experiences and informs more holistic care approaches. These chapters serve as the culmination of the thesis, demonstrating the tangible impact that enhanced transparency can have on real-world medical practices, particularly in oncology and personalised medicine.

Transparency policies will only succeed if they provide clear benefits to all parties involved. For pharmaceutical companies, this could mean greater trust and acceptance of new therapies by the public. For researchers, streamlined access to data will enable more robust analyses and innovative discoveries. For patients, data transparency translates into better care and more reliable treatments.

Therefore, this thesis argues that transparency should not just be a bureaucratic goal - but a moral imperative. The chapters explore how reforming data sharing practices can break down the walls of isolation and secrecy that impede scientific progress. The future of medicine will be defined not by the accumulation of isolated datasets, but by the shared experiences, collaborations, and insights they inspire. In this, my work aims to make a case for nothing short of a revolution in clinical trial transparency - one that demands collective action and promises transformative outcomes. The practical applications chapters chart this path where they offer a blueprint for how we can turn clinical trial data into a living, breathing tool for discovery.

Table 1: Thesis outline

	Chapter	Chapter Title	Presented Publications
	1	Introduction	
Section 1: Transparency of industry-sponsored clinical trial data	2	Audit of Data Sharing by Pharmaceutical Companies for Anticancer Medicines Approved by the US Food and Drug Administration	Modi ND, Abuhelwa AY, McKinnon RA, et al. Audit of Data Sharing by Pharmaceutical Companies for Anticancer Medicines Approved by the US Food and Drug Administration. *JAMA Oncol.* 2022; 8(9):1310–1316
	3	Clinical Study Report and Individual Participant Data Transparency for FDA-Approved Anticancer Drugs: A Call for Systematic Data Availability	Modi ND, Swain SM, Buyse M, et al. Clinical Study Report and Individual Participant Data Transparency for FDA-Approved Anticancer Drugs: A Call for Systematic Data Availability. *Journal of Clinical Oncology. 2024: 1-5*
	4	The State of Individual Participant Data Sharing for Highest Revenue Medicines	Modi ND, Li LX, Logan JM, et al. The state of individual participant data sharing for the highest-revenue medicines. *Clinical Trials. 2024; 1-8.*
	5	A 10-year update to the principles for clinical trial data sharing by pharmaceutical companies: perspectives based on a decade of literature and policies	Modi ND, Kichenadasse G, Hoffman TC, et al. A 10-year update to the principles for clinical trial data sharing by pharmaceutical companies: perspectives based on a decade of literature and policies. *BMC Med. 2023; 21, 400*
Section 2: Practical Applications of Enhanced Data Transparency	6	Prediction of severe neutropenia and diarrhoea in breast cancer patients treated with abemaciclib	Modi ND, Abuhelwa AY, Badaoui S, et al. Prediction of severe neutropenia and diarrhoea in breast cancer patients treated with abemaciclib. *The Breast.* 2021; 58:57–62
	7	Patient-reported outcomes predict survival and adverse events following anticancer treatment initiation in advanced HER2-positive breast cancer	Modi ND, Danell NO, Perry RNA, et al. Patient-reported outcomes predict survival and adverse events following anticancer treatment initiation in advanced HER2-positive breast cancer. *ESMO Open.* 2022; 7(3):100475
	8	Significance, Future Directions and Conclusion	
	9	Bibliography	
	10	Appendices	

CHAPTER 2: AN AUDIT OF DATA SHARING BY PHARMACEUTICAL COMPANIES FOR ANTICANCER MEDICINES APPROVED BY THE FDA

Note: This chapter is adapted from a published manuscript. Modi ND, Abuhelwa AY, McKinnon RA, Boddy AV, Haseloff M, Wiese MD, Hoffmann TC, Perakslis ED, Rowland A, Sorich MJ & Hopkins AM. Audit of data sharing by pharmaceutical companies for anticancer medicines approved by the US Food and Drug Administration. JAMA Oncology. 2022; 8: 1310-1316. DOI: https://doi.org/10.001/jamaoncol.2022.2867

Contributions: In this publication, I, Modi ND, contributed significantly to the research design, data collection and analysis, and the writing and editing of the manuscript. Specifically, I was responsible for 70% of the research design, 90% of the data collection and analysis, and 90% of the writing and editing. Hopkins AM contributed to 10% of the research design, 10% of the data collection and analysis, and 1% of the writing and editing. Abuhelwa AY, McKinnon RA, Boddy AV, Haseloff M, Wiese MD, Hoffmann TC, Perakslis ED, Rowland A and Sorich MJ contributed the remaining portions of the research design, data collection and analysis, and writing and editing in varying percentages.

Introduction

Decisions by regulators and clinicians on whether to approve and use new medications are typically based on findings from pivotal clinical trials [22]. For most newer medicines an industry sponsor drives the early generation of the evidence base supporting the medicine, but this requires access to and facilitation by global health care systems [23]. Data from early clinical trials remain the centrepiece of safety and efficacy assessments, at least until post-marketing data can reach maturity [1, 24]. Transparent sharing of IPD from clinical trials facilitates enrichment of the post-approval evidence-base through novel secondary analyses and informs the design of future studies [2, 22, 24-29].

In 2010, the European Medicines Agency (EMA) began adopting forward-looking policies promoting clinical trial data sharing upon market authorisation [22]. In 2013, the pharmaceutical industry, via the Pharmaceutical Research and Manufacturers of America (PhRMA) and the European Federation of Pharmaceutical Industries and Associations (EFPIA), acknowledged the importance of IPD sharing and endorsed a commitment to sharing anonymised IPD for approved medicines upon request by qualified researchers [13, 30]. A 2018 audit reported that only 15% of clinical trials were available for sharing 2 years post-publication of the primary outcome, with no sharing occurring for oncology trials [11]. Since 2018, there has been significant development of resources and systems to facilitate research using transparently shared IPD [15, 16, 25, 31-34], and progress has been made by the pharmaceutical industry to develop data sharing policies and processes. Thus, the status of IPD sharing of pivotal oncology trials warrants re-examination.

Sharing of IPD is important for highly utilised newer medicines where the evidence underpinning use is almost exclusively derived from clinical trials supporting the medicine approval. Despite being one of the most active areas for drug development over the last decade, there is limited data regarding the sharing of anonymised IPD underlying pivotal oncology trials for newer anticancer medicines. This study evaluated the eligibility of independent qualified researchers to access IPD from oncology trials that supported the United States Food and Drug Administration (FDA) approval of new anti-cancer medicines within the last 10 years.

Methods

Sample and data

A structured search was undertaken to identify all anticancer medicines approved by the FDA between 1 Jan 2011 to 30 Jun 2021 [35, 36]. For these anticancer medicines, product labels were accessed through the US National Institute of Health website [37] and a list of the clinical trials that had their results summarised in the product labels was made. For each trial, information on the National Clinical Trial (NCT) number, phase of the trial, the date the trial results were added to the product label, and the owner of the medicine was collected. Information on the primary sponsor, cancer type (solid or haematological), start date, primary completion date, and final completion date was collected from ClinicalTrials.gov. For industry-sponsored trials, data were collated on whether the sponsoring pharmaceutical company (i.e., the owner of the medicine) was within the top 20 by global revenue [38, 39], and whether the investigated medicines were within the top 10 anticancer medicines by global sales [40, 41]. The PhRMA and/or EFPIA membership statuses of trial sponsors were documented, and websites of trial sponsors were searched to identify the presence of a public IPD sharing policy. The data sharing policies were utilised to collate information on the data sharing process (i.e., a company internal or company external process [e.g., vivli.org (Vivli), clinicalstudydatarequest.com (CSDR), or yoda.yale.edu (YODA)]), contact details for IPD sharing enquiries, and whether trial completion was a criterion for data sharing (i.e., did the policy contain a statement that IPD would not be available until after cessation of follow-up data collection).

Determination of IPD sharing eligibility

Beginning 1 Aug 2021, the IPD sharing eligibility of each trial was confirmed by either identification of a public listing of the trial as eligible for IPD sharing or receipt of a positive response to a standardised enquiry (Appendix 1) directed to the trial sponsor (or medicine owner if different). Ineligibility for IPD sharing was confirmed by a negative response to the enquiry (i.e., receipt of written confirmation from the trial sponsor that IPD would not be shared with independent researchers). If a trial was indicated as not eligible for IPD sharing, details of the reason(s) for

ineligibility and when the trial would become eligible were requested. If no response to the initial enquiry was received, prompts were sent 30 and 60 days after the initial enquiry. If no response had been received by the trial sponsor (and medicine owner if different) by 120 days from the initial enquiry, then the trial was deemed to be ineligible for IPD sharing.

Analysis

All statistical analyses were performed using R 4.1.2 [42]. Forest plots and Chi-square tests were used to evaluate and present differences in trial IPD sharing eligibility proportions according to key descriptive company, drug, and trial-level subgroups.

Ethics

The research undertaken was assessed as negligible risk research and was confirmed exempt from requiring Flinders University Human Research Ethics Committee review.

Results

Sample

Over the 10-year sampling period, 115 anticancer medicines were approved by the FDA – 84% (n=96) of them are also presently approved by the EMA. These medicines were owned/co-owned by 49 pharmaceutical companies, and their approval was based on the results of 304 industry-sponsored trials. All trials were registered on clinicaltrials.gov. Of the 304 trials, 16 evaluated cytotoxic medicines, 12 hormonal medicines, 80 immunomodulators, and 196 targeted therapeutics not elsewhere specified. Table 2 provides a detailed summary of trials by cancer subtype.

Table 2: Summary of trials by cancer subtype

Cancer subtype	Number of trials
Lung Cancer	39
Leukemia	28
Lymphoma	27
Breast Cancer	26
Myeloma	25
Melanoma	21
Prostate Cancer	15
Bladder Cancer	9
Liver Cancer	8
Colon and Rectal Cancer	8
Kidney Cancer	8
Ovarian Cancer	6
Oesophageal Cancer	4
Stomach Cancer	3
Other Solid Cancers	56
Other Haematological Cancers	21

Of the 304 trials, 203 (67%) were in patients with solid tumours and 101 (33%) in patients with haematological malignancies. There were 199 (65%) Randomised trials and 105 (35%) non-randomised trials, including 16 (5%) phase 1, 112 (37%) phase 2, and 176 (58%) phase 3 trials. Of the 304 trials, 140 (46%) had a trial start date before 1 Jan 2014, whereas 164 (54%) had a trial start date after 1 Jan 2014. Less than 3 years had passed since the result summary was added to the product label for 136 (45%) trials, 3 to 7 years for 126 (41%), and more than 7 years for 42 (14%) trials.

For the 49 pharmaceutical companies audited, 24 were PhRMA/EFPIA members and 28 had a publicly available IPD sharing policy; these companies sponsored 261 (86%) and 273 (90%) of the trials audited, respectively. Nineteen pharmaceutical companies share IPD via an external platform (i.e., Vivli (n=13), CSDR (n=5), and YODA (n=1)), 9 share via a company internal process, and 21 had no defined process to share IPD; these companies sponsored 211 (70%), 62 (20%), and 31 (10%) of the trials included, respectively.

Eighteen of the top 20 pharmaceutical companies by global revenue sponsored trials in the study sample. These 18 companies owned 81 of the audited medicines for which results from 245 (81%) trials were summarised in their product labels. Additionally, for the top 10 anticancer medicines by global revenue, results from 89 (29%) trials were summarised in their respective product labels.

Of the 304 industry-sponsored trials audited, the eligibility for IPD sharing status was publicly available for 64 (21%) trials. The remaining 240 (79%) trials required an enquiry to the sponsor to establish if the trial was eligible for data sharing. The median (IQR) response time to these enquiries was 42 days (7–60 days). For 9 trials, sponsored by 8 different pharmaceutical companies, no response to the eligibility enquiries were received.

Eligibility to Share

Of the 304 included trials, 136 (45%) were indicated as eligible for IPD sharing with independent researchers.

Figure 1 presents trial IPD sharing eligibility for pharmaceutical companies within the top 20 by global revenue and top 10 anticancer medicines by global revenue. Of the top 20 pharmaceutical companies by revenue, four companies (AbbVie, Bayer, Gilead Sciences, and Takeda) had less than 50% of their oncology trials available for IPD sharing, and five companies (Astellas, Bristol Myers Squibb, GlaxoSmithKline, Merck Sharp & Dohme Corp, and Teva) had less than 10% available for IPD sharing. Of the top 10 anticancer medicines by global revenue, less than 10% of trials on nivolumab, pembrolizumab, and pomalidomide were available for IPD sharing.

Figure 1: Eligibility to share by A) Top 10 anticancer medicines B) Top 20 Pharmaceutical companies by revenue.

n: number of trials evaluated.

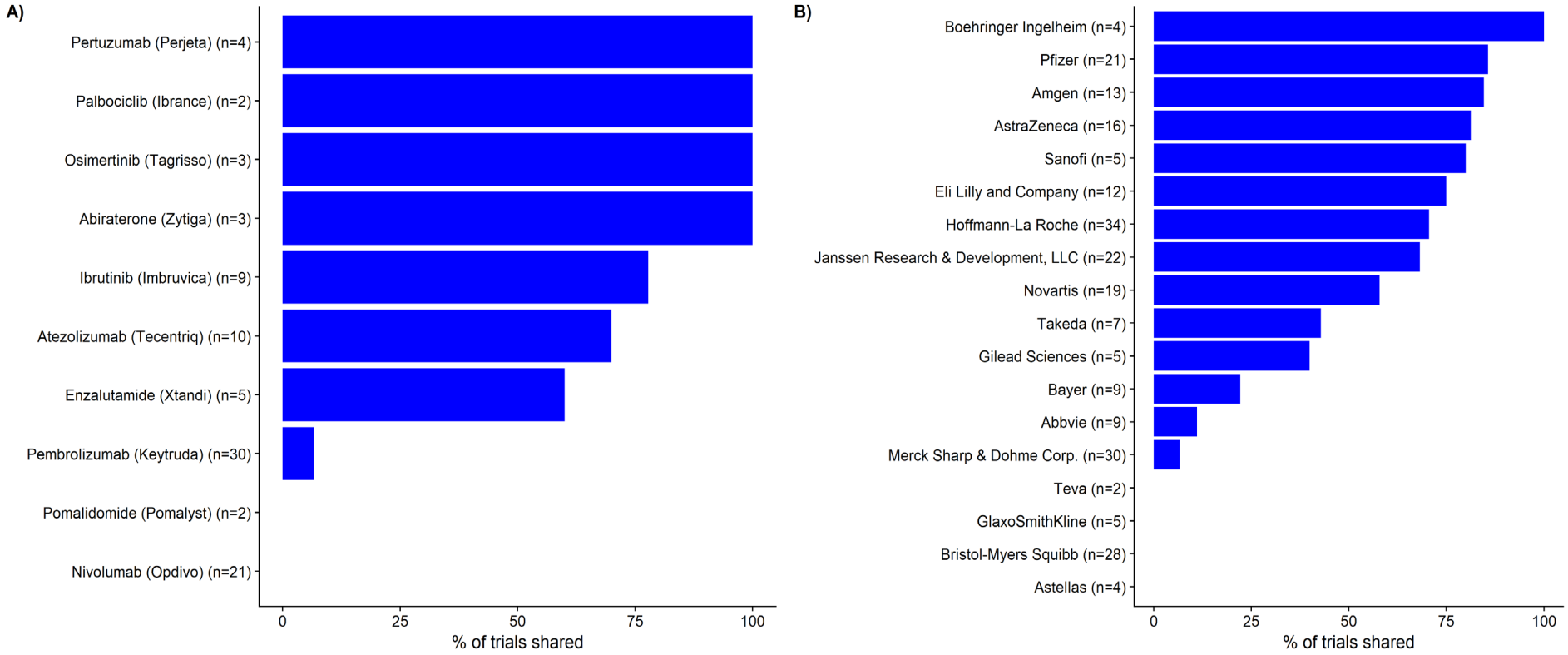


Figure 2 and Table 3 present the proportion of trials eligible for IPD sharing according to key descriptive subgroups. Being a PhRMA/EFPIA member (48% vs. 26%, P<0.01) and a company with a publicly available IPD sharing policy (49% vs. 10%, P<0.01) was associated with a higher proportion of trials being eligible for IPD sharing. Pharmaceutical companies within the top 20 by global revenue also had a significantly higher proportion of trials eligible for IPD sharing compared to companies outside the top 20 (49% vs. 29%; P<0.01), and companies that shared IPD on an external platform had a higher proportion of trials eligible for IPD sharing than companies that shared via an internal process or had no formal process to share IPD (54% vs. 32% vs. 10%, respectively P<0.01).

Figure 2: Breakdown of IPD sharing eligibility according to key descriptive subgroups

Category	No. of trials	No. of companies	No. eligible to share IPD (%, 95% CI)	*P* value
PhRMA/EFPIA membership				
Member	261	24	125 (48, 42-54)	.006
Nonmember	43	25	11 (26, 13-41)	
Cancer type				
Solid	203	35	89 (44, 37-51)	.66
Hematologic	101	32	47 (47, 37-57)	
Trial phase				
Phase 1	16	12	2 (13, 2-38)	
Phase 2	112	12	43 (38, 29-48)	.002
Phase 3	176	34	91 (52, 44-59)	
Trial design				
Nonrandomized	105	38	35 (33, 24-43)	.003
Randomized	199	33	101 (51, 44-58)	.006
Company within the top 20 by global revenue				
Top 20	245	18	119 (49, 42-55)	.02
Not top 20	59	31	17 (29, 18-42)	
Anticancer medicine within the top 10 by global sales				
Top 10	89	9	31 (35, 25-46)	<.001
Not top 10	215	49	105 (49, 42-56)	
Process of sharing				
External	211	19	113 (54, 46-60)	<.001
Internal	62	9	20 (32, 21-45)	
None identified	31	21	3 (10, 2-26)	
Public data sharing policy				
Available	273	28	133 (49, 43-55)	<.001
Not available	31	21	3 (10, 2-25)	
Trial start date				
Before January 1, 2014	140	30	88 (63, 54-71)	<.001
On or after January 1, 2014	164	42	48 (29, 22-37)	
Time since trial listed in product label				
<3 y	136	44	40 (29, 22-38)	<.001
3-7 y	126	22	65 (52, 43-61)	
>7 y	42	17	31 (74, 58-86)	
IPD sharing policy includes a criterion about trial completion				
No	108	15	62 (57, 47-67)	<.001
Yes	165	13	71 (43, 35-51)	
Has no policy	31	21	3 (10, 2-26)	

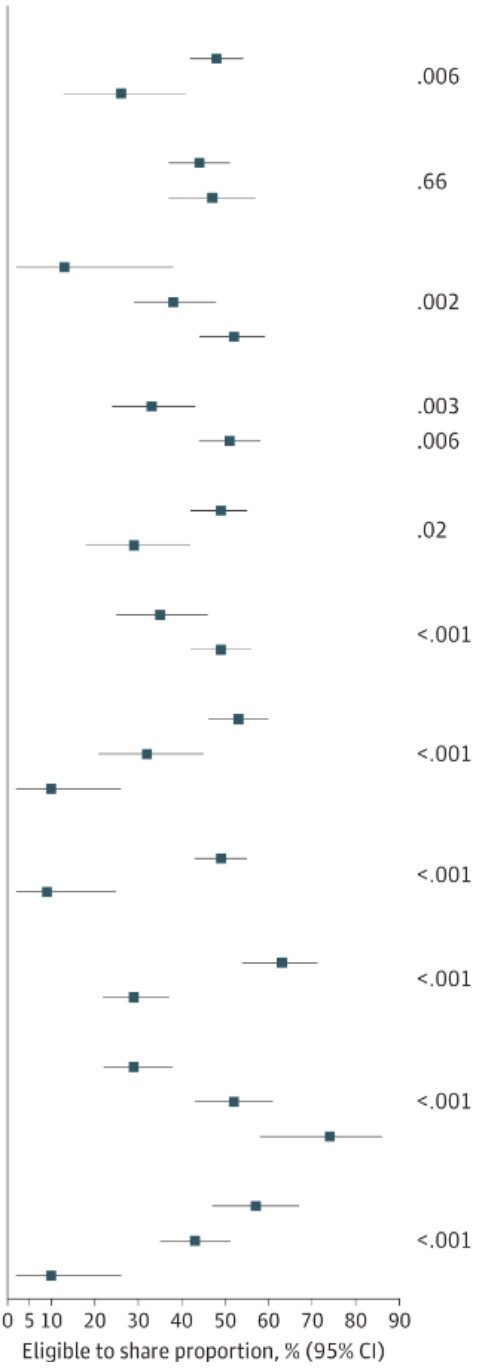


Table 3: Breakdown of IPD sharing eligibility according to key descriptive subgroups

Category		Companies (n)	Trial eligible for IPD sharing		P-value*
			Yes	No	
PhRMA/EFPIA membership	Member	24	125 (48%)	136 (52%)	<0.01
	Non-Member	25	11 (26%)	32 (74%)	
Cancer Type	Solid	35	89 (44%)	114 (56%)	0.66
	Haematological	32	47 (47%)	54 (53%)	
Trial Phase	Phase 1	12	2 (13%)	14 (87%)	<0.01
	Phase 2	12	43 (38%)	69 (62%)	
	Phase 3	34	91 (51%)	85 (49%)	
Trial Design	Non-Randomised	38	35 (33%)	70 (67%)	<0.01
	Randomised	33	101 (51%)	98 (49%)	
Company within the top 20 by global revenue	Top 20	18	119 (49%)	126 (51%)	<0.01
	Not Top 20	31	17 (29%)	42 (71%)	
Top 10 anticancer medicine by global sales	Top 10	9	31 (35%)	58 (65%)	0.02
	Not Top 10	49	105 (49%)	110 (51%)	
Process of sharing	External	19	113 (54%)	98 (46%)	<0.01
	Internal	9	20 (32%)	42 (68%)	
	None Identified	21	3 (10%)	28 (90%)	
Public data sharing policy	Available	28	133 (49%)	140 (51%)	<0.01
	Not Available	21	3 (10%)	28 (90%)	
Trial Start Date	< 1 Jan 2014	30	88 (63%)	52 (37%)	<0.01
	≥ 1 Jan 2014	42	48 (29%)	116 (71%)	
Time since trial listed in product label	<3 years	44	40 (29%)	96 (71%)	<0.01
	3-7 years	22	65 (52%)	61 (48%)	
	>7 years	17	31 (74%)	11 (26%)	
IPD sharing policy includes a criterion about trial completion	No	15	62 (57%)	46 (43%)	<0.01
	Yes	13	71 (43%)	94 (57%)	
	Has no policy	21	3 (10%)	28 (90%)	

Chi-sq evaluation of the distribution between trial eligible for IPD sharing "Yes" and "No"

In contrast, the proportion of trials eligible for IPD sharing was significantly lower for the top 10 anticancer medicines by global sales as compared to anticancer medicines, not in the top 10 (35% vs. 49%; P=0.02). The eligibility proportion was lower for non-randomised trials compared to Randomised trials (33% vs. 51%, P<0.01), phase 1 compared to phase 2 and phase 3 trials (13% vs. 38% vs. 51%, respectively P<0.01), and for trials with a trial start date after 1 Jan 2014 (29% vs. 63%, P<0.01). Eligibility was also lower for trials listed in a product label less than 3 years ago as compared to trials listed in a product label 3-7 years ago or more than 7 years ago (29% vs. 52% vs. 74%, respectively P<0.01). Finally, the proportion of trials eligible for IPD sharing was lower when the sponsoring company had an IPD sharing policy that included a criterion about trial completion as compared to companies without such a criterion (43% vs. 57%, P=0.02). No difference in IPD sharing eligibility was observed between trials for anticancer medicines used to treat solid tumours compared to those used for haematological malignancies.

Data were confirmed as not available for sharing for 168 trials (55%) (Table 4). The most common reason communicated for trial IPD sharing ineligibility was that the trial was still ongoing (89 trials [53%]) - i.e., the sponsor indicated that follow-up was continuing for the trial and as such the IPD for the results reported in the product label were not available for sharing (Table 4). On clinicaltrials.gov the final completion dates for these 89 trials ranged from 2020 to 2027, with 59 (66%) documented as having passed primary completion and 9 (10%) as fully complete. A further 21 (12%) were indicated as ineligible for IPD sharing because, despite passing their final completion date, the IPD remained under embargo. Table 3 documents the reasons provided for trial IPD sharing ineligibility.

Table 4: Breakdown of reasons for IPD sharing ineligibility provided by the clinical trial sponsor

Reason	Total cohort (n=304)	Trials by company within the top 20 by global revenue (n=245)	Trials for top 10 anticancer medicines by global sales (n=89)
Number of trials confirmed as ineligible for IPD sharing	**168**	**126**	**58**
• Study still ongoing	89	75	37
• Study has passed final completion, but IPD still under embargo	21	17	14
• Medicine not approved by both EMA and FDA, or ongoing regulatory submission	12	9	2
• Phase I/II trials are out of scope	10	9	0
• Consent Form Issues	9	5	3
• Sponsor does not share IPD	6	0	0
• No response received within 4 months	9	0	0
• Other	12	11	0

Due to infrequency in the sample, Appendix 2 summarises the IPD sharing eligibility of non-industry sponsored trials. The raw dataset generated and analysed in this study can be accessed online [43].

Discussion

To our knowledge, this is the largest structured study to assess IPD sharing eligibility of clinical trials for recently approved medicines, and it is the first to evaluate data sharing for pivotal industry-sponsored oncology trials. In our sample of 304 trials underpinning the FDA approval of 115 new anticancer medicines over the past 10 years, 45% of those trials were confirmed eligible for IPD sharing. However, 55% of the queried trials were confirmed as not available for IPD sharing. With profit correlating to global drug utilisation, it is a missed opportunity that Astellas, Bristol Myers Squibb, GlaxoSmithKline, Merck Sharp & Dohme Corp, and Teva had less than 10% of their sampled oncology trials available for IPD sharing.

Prior to this study, the largest structured assessments of the eligibility of independent researchers to request industry-sponsored clinical trial IPD were conducted by Murugiah et al [44] and Hopkins et al [11]. In 2016 Murugiah et al [44] identified that IPD was eligible for sharing from approximately 25% of large cardiovascular trials, while in 2018 Hopkins et al [11] documented that only 15% of

clinical trials were available for IPD sharing 2 years post-publication of primary results, with no sharing occurring for the oncology trials in the sample. The present study demonstrates a significant increase in IPD sharing for major oncology trials, with 45% (136 of 304) of the audited trials confirmed as eligible. Further exemplifying improvements in IPD sharing awareness by the pharmaceutical industry is that in 2018 Hopkins et al [11] did not receive a response to 36% of IPD sharing enquires. In the present study, the no response rate to IPD sharing enquiries was only 3%. For those 9 trials, sponsored by 8 different pharmaceutical companies, where no response to the eligibility enquiries was received, it may be worth considering whether the country in which these companies are based plays a role in their response rates. However, in this case, six companies were based in the US, one in the EU, and one in China, making it unlikely that any meaningful patterns could be drawn from this small, geographically skewed sample. Nevertheless, evaluating regional trends in IPD sharing responses may be an area for future exploration with a larger, more diverse dataset. Herein, only 32% of evaluated pharmaceutical companies outside the top 20 by global revenue had a publicly accessible data sharing policy, in stark contrast to the 100% of sponsors within the top 20 by global revenue. These findings are in line with prior research [11, 44-46], and demonstrate a continued need for advocacy and support to smaller pharmaceutical companies' to enable data transparency – as, despite their lower revenue, they anchor the registration of a significant portion of critical innovator medicines in oncology.

In this study, we confirmed that IPD was eligible for sharing from 136 industry-sponsored trials that had results summarised in the product labels of 60 anticancer medicines approved by the FDA over the last 10 years. These trials consist of over 70,000 patients and provide an immense opportunity for independent scientific investigations by regulators, clinicians, and researchers. Notably, this includes IPD sharing for over 50% of the trials summarised in the product labels of atezolizumab, abiraterone, enzalutamide, ibrutinib, osimertinib, palbociclib, and pertuzumab. Further, five of the top 20 pharmaceutical companies by revenue (Amgen, AstraZeneca, Boehringer Ingelheim, Pfizer, and Sanofi) indicated that more than 75% of their sampled oncology trials were eligible for IPD sharing. These are noteworthy achievements for transparency catalysed by the 2014 PhRMA/EFPIA guiding principles on responsible data sharing [13, 30], and it provides a beacon of hope to one of the least trusted industries in the world [13, 24].

On the other hand, IPD was unavailable for sharing from 168 investigated industry-sponsored trials that had results summarised in the product labels of 78 anticancer medicines approved by the FDA in the last 10 years. These trials included over 85,000 participants. It is morally indefensible that this wealth of anonymised IPD remains unavailable to independent investigation, despite the rollout of the medicines within the US and global cancer populations based on the product label results. When

participants commit to these trials, they are generally advised and reasonably expect that although they may not personally benefit from their participation, the knowledge gained will contribute to better care for future patients. Clearly, this commitment to participants in oncology trials is not yet being fully met. Several potential reasons may contribute to this issue. Some pharmaceutical companies may be reluctant to share IPD due to concerns about competitive advantage, proprietary knowledge, or unresolved legal or regulatory barriers. For instance, trials involving collaboration agreements, such as the one between Takeda and Seagen, are often subject to complex contractual arrangements that delay or restrict data sharing. Additionally, trials completed before 2014 were excluded from sharing due to falling outside of data sharing policies, reflecting outdated or arbitrary policy cut-offs that do not account for the relevance of older trials. Trials involving rare conditions, where data sharing could be complicated by small sample sizes or confidentiality risks, further illustrate these barriers. The “Other” category in Table 4 highlights these additional complexities.

A significant strength of this study is that it is the first to have the sample size to facilitate key company, drug, and trial-level subgroup evaluations. Herein, trials were more likely to be eligible for IPD sharing if the medicine owner was a PhRMA/EFPIA member, had a publicly available IPD sharing policy, shared data via an external platform, and were within the top 20 by global revenue. Randomised and phase 3 trials were also more likely to be eligible for IPD sharing. Trials on top 10 anticancer medicine by global sales, that were more recently initiated or listed in a product label, and those performed by a sponsor with an IPD sharing policy including a criterion based on trial completion were less likely to be eligible for IPD sharing. Based on these findings, to improve data sharing it is recommended that medicine sponsors (1) join PhRMA/EFPIA and establish a data sharing policy, (2) establish IPD sharing processes external to the company (i.e., recognising that there is an internal conflict of interest), and (3) have a policy that states all IPD underlying results presented in a product label will be immediately eligible for sharing. The third recommendation aims to remove overly long embargo periods and ensure that registration of a medicine, an event that allows widespread use, immediately triggers sharing of the clinical trial data supporting the registration. However, long-term follow-up of clinical trials remains essential and sharing of IPD following registration should not jeopardise efforts to collect data on long-term outcomes. Noticeably, the reason for the unavailability of 89 of the 168 trials ineligible for IPD sharing was that the trial was still ongoing, and it is concerning that by our estimate, 50% of these trials will still be unavailable for sharing in two years. Highlighting the importance of policies that facilitate the sharing of primary outcome data, 109 of the 115 anticancer medicines evaluated herein at some point had the results of a trial summarised in the product label before passing the follow-up completion date on clinicaltrials.gov.

A recent editorial in the British Medical Journal (BMJ) called for pharmaceutical companies to update their policies to facilitate IPD sharing for newly registered medicines [24]. The editors emphasised that due to embargo criterion within the transparency policies of Pfizer and Moderna, IPD underpinning the registration of COVID-19 vaccines will not be available for years – which is not acceptable given the scale and current relevance of their use. Herein, we make a similar call to the justifiability of 90% of the trials summarised in the product labels of nivolumab, pembrolizumab, and pomalidomide being ineligible for IPD sharing. These drugs represent major health initiatives reaping vast profits; therefore, patients deserve the confidence that all opportunities to understand the benefits and harms of the treatment for their circumstances have been made and that the scientific claims have been independently scrutinised. This is currently not the case.

A study limitation is that our sample was focused on anticancer medicines registered by the FDA over the past 10 years. The time was chosen as the product label results are the centrepiece of safety and efficacy for newer medicines, and correspondingly the data should be subject to high scrutiny. However, the generalisability of the findings to older medicines or medicines solely approved by the EMA is unknown. Further, as most newer medicines have been developed by an industry sponsor, the sample of non-industry sponsored trials was too small for reliable comparison. Inherently, the data sharing practices of non-industry trial sponsors (e.g., academic and health institutions) deserve purposeful audit. Finally, future studies should investigate the time from proposal submission to data receipt (time to access should approximate 4 months) [16, 31], data completeness upon receipt, researcher support initiatives (e.g., data dictionaries) [29], and the sharing of data for medicines beyond cancer treatment. It is noteworthy that during the 10-year sampling period the FDA approved 437 novel therapies.

In conclusion, the present study found that 136 (45%) trials underpinning the FDA approval of 115 anticancer medicines over the past 10 years were indicated as eligible for IPD sharing. This demonstrates a significant increase in IPD sharing for industry-sponsored oncology trials over the past 5-years and represents a significant resource for scientific discovery. Nonetheless, 55% of the queried trials were confirmed as not available for IPD sharing, and, notably, less than 10% of the trials with results summarised in the product labels of nivolumab, pembrolizumab, and pomalidomide were available. Since these trials form the basis of these newer medicines' safety and efficacy claims, we question whether it is justified that the data is unavailable to independent scrutiny. The present study's findings reiterate calls that transparency policies need updating so that all IPD informing results presented in a product label or underpinning drug registration are immediately eligible for sharing.

CHAPTER 3: CLINICAL STUDY REPORT AND INDIVIDUAL PARTICIPANT DATA TRANSPARENCY FOR FDA-APPROVED ANTICANCER DRUGS: A CALL FOR SYSTEMATIC DATA AVAILABILITY

Note: This chapter is adapted from a published manuscript. Modi ND, Swain SM, Buyse M, Kuderer NM, Rowland A, Rockhold FW, Sorich MJ & Hopkins AM. Clinical Study Report and Individual Participant Data Transparency for FDA-Approved Anticancer Drugs: A Call for Systematic Data Availability. Journal of Clinical Oncology. 2024. DOI: https://doi.org/10.1200/JCO.24.00539

Contributions: In this publication, I, Modi ND, contributed significantly to the writing and editing of the manuscript. Specifically, I was responsible for 90% of the writing and editing. Swain SM, Buyse M, Kuderer NM, Rowland A, Rockhold FW, Sorich MJ & Hopkins AM contributed the remaining portions of the writing and editing in varying percentages.

Introduction

In the previous chapter, we conducted the largest structured study to date on IPD sharing eligibility of clinical trials for newly approved anticancer medicines. We found that while 45% of the trials underpinning FDA approvals were eligible for IPD sharing, a significant 55% remained inaccessible, representing a missed opportunity for advancing scientific discovery. This chapter builds on those findings to emphasise the broader necessity of making both CSRs and IPD systematically available to enhance transparency and scientific discovery in oncology.

Organizations including the World Health Organization (WHO), Cochrane, International Committee of Medical Journal Editors (ICMJE), EMA and patient advocacy groups have all argued that providing independent access to CSRs and IPD from clinical trials is important for building trust in drug approval processes, preventing study duplications, and informing the design of future trials [2, 47-49]. CSRs, most often associated with industry-sponsored clinical trials, are comprehensive documents that frequently span hundreds of pages to provide detailed aggregate-level insights into clinical trial methodologies and results. IPD, on the other hand, offers in-depth information on each trial participant's demographics, laboratory measurements, adverse events, and responses to treatment, as recorded during each clinical visit. The comprehensive sharing of CSRs and IPD has immense potential to enable the discovery of novel insights through new subgroup and secondary endpoint analyses, support meta-analyses, and foster a deeper understanding of drug effects [2, 14]. These are particularly important initiatives for adverse event information on newer drugs and questions that single Randomised controlled trials (RCTs) are not statistically powered to answer. Moreover, it should be acknowledged that participants often enrol in clinical trials with the understanding that their involvement, while potentially not personally beneficial, will contribute to

advancing future care [26]. The responsibility thus falls on all stakeholders to honour this commitment by maximising the potential for scientific discovery.

In this commentary, we aim to emphasise that high-quality systematic reviews and meta-analyses play an important role in informing clinical practice guidelines and facilitating personalised clinical care. Traditionally, such reviews utilise summary data extracted from publications and abstracts. This approach is inherently constrained by the limitations of data reported within these documents. Greater sharing and accessibility of CSRs and IPD substantially expands the scope of important clinical questions that can be investigated. For instance, the importance of IPD was demonstrated in a recent study clarifying survival outcomes for patients with prostate cancer over 80 years of age treated with androgen receptor inhibitors [50]. Similarly, IPD meta-analyses in breast cancer have led to practice-changing insights, such as identifying survival benefits from regional node radiotherapy and underscoring the benefits of aromatase inhibitors in women with estrogen receptor-positive disease [51, 52]. Furthermore, it is widely acknowledged that information on adverse events in journal articles is less detailed than in CSRs, underscoring the importance of CSRs for thorough safety assessments.

However, when conducting systematic reviews and meta-analyses with the intention to use IPD or CSRs, it must be recognised that omitting trials because these data sources are unavailable introduces a substantial risk of bias [53]. The more trials with unavailable data, the greater the risk of compromising the study's validity and ability to influence and improve clinical care [53]. To minimise the risk of bias, it is recommended to achieve an inclusion rate of at least 90% of relevant participants/trials [54]. In the context that comprehensive access to data can facilitate practice-changing findings, we explore the recent improvements, current state, and future opportunities for CSR and IPD sharing concerning anticancer drugs. Recognising that currently, the use of IPD or CSRs is not the norm in systematic reviews and meta-analyses.

Recent Policy Changes in CSR and IPD Sharing

Historically, the sharing of CSRs and IPD by trial sponsors, both in academia and industry, has been limited. However, the last decade has seen growing consensus within industry, academia, and regulatory bodies towards improving clinical trial transparency [13, 55-57]. While this article focuses on industry-sponsored trials - the primary evidence base for newly approved drugs - the strategies for improving data sharing practices are equally applicable to non-industry/academic trial sponsors, where data sharing also often falls short. CSRs are key documents in the drug approval process, affirming the safety and efficacy of drugs by providing detailed information on trial methodologies and results, including specifics often not found in journal publications or summary reports. The EMA

in 2015 [55] and Health Canada in 2019 [57] implemented policies and online platforms to facilitate the public sharing of CSRs submitted in drug approval applications. The U.S. FDA has also expressed support for CSR sharing [56], although it advocates for sponsor-led libraries to reduce administrative burdens, without implementing a policy to directly facilitate CSR sharing themselves. Admirably, many pharmaceutical companies have established data sharing policies, promoting independent researchers can request access to their CSRs [13, 14].

IPD encompasses the demographic, clinical measurement, and outcome data collected on each participant in a clinical trial. Responding to growing calls for transparency, since 2013, most large pharmaceutical companies have established data sharing policies that allow independent researchers to request access to de-identified IPD from their clinical trials [13, 14, 18, 43]. This has seen the establishment of data sharing platforms, like ClinicalStudyDataRequest.com, Yale University Open Data Access Project, Vivli, and Project Data Sphere, significantly supported by the pharmaceutical industry. Additionally, a growing number of non-industry trial sponsors now require the inclusion of data sharing plans in research grant applications [58].

Current Status of CSR and IPD Sharing for Anticancer Drugs

To understand the impact of recent policy changes, a comprehensive cross-sectional analysis was undertaken of 114 industry-sponsored oncology trials, which contributed to the FDA-authorised drug labels of 2021's top 10 highest-revenue anticancer medicines [59]. This study revealed that 35% of the trial CSRs were available for download via EMA/Health Canada portals, with an additional 21% eligible for request from the sponsor [59]. The availability of 56% of CSRs marks significant progress over the past decade, although the unavailability of 44% of CSRs poses challenges for conducting valid systematic reviews requiring such data [53].

Regarding IPD accessibility, a recent JAMA Oncology publication evaluating 304 industry-sponsored oncology trials, linked to the FDA-authorised drug labels of anticancer medicines approved from 2011 to 2021, identified that 45% of these trials were eligible for IPD request [43]. Subsequently, when IPD was formally requested, data were only provided for 77% of the solid-tumour trials previously confirmed to be eligible for request [18]. The median time to data provision was 123 days (range, 117-352 days) with substantial variability observed in the completeness of provided data [18].

Overall, IPD was acquired from 34% of trials linked to FDA-authorised drug labels for solid tumour treatments approved between 2011 and 2021 [18, 43], marking a major improvement compared to what was possible 10 years ago. However, relying on IPD from only a third of relevant trials to

conduct systematic reviews introduces a high risk of bias [53]. Consequently, the current ability to use shared trial IPD to efficiently drive meaningful improvements in clinical practice and patient outcomes is still limited.

How to achieve comprehensive CSR sharing?

Addressing the challenges of data sharing requires a multifaceted approach that engages all stakeholders, including trial participants, funding bodies, journals, regulators, professional organizations, and the pharmaceutical industry (Table 5).

Table 5: Current Accessibility and Proposed Strategies to Improve the Sharing of Clinical Study Reports and Individual Participant Data from Industry-Sponsored Oncology Trials

Aspect	Clinical Study Reports (CSRs)	Individual Participant Data (IPD)
Definition	• Comprehensive documents prepared by pharmaceutical companies to provide in-depth detail on trial methodologies and results.	• The detailed data collected on each participant in a clinical trial, including the demographic, clinical measurement and outcomes information.
Key Benefits of Sharing	• Enhances transparency to trial outcomes supporting regulatory decisions. • Provides access to detailed safety and efficacy information that can facilitate meta-analyses and enhance clinical knowledge.	• Enables novel subgroup analyses, comprehensive IPD meta-analyses and helps inform the design of future trials. • Can help minimise the occurrence of redundant research, build trust in the pharmaceutical industry and inform precision medicine approaches.
State of Data Accessibility	• Approximately 35% of the CSRs supporting the approval of the top 10 highest revenue anticancer medicines are available for download from EMA/Health Canada. • An additional 21% are eligible for request from the trial sponsor, however, such processes may take time. • Over-redactions of publicly accessible CSRs have been observed.	• Approximately 45% of the IPD from trials underpinning the approval of anticancer medicines in the last decade are eligible for request by independent researchers. • While trials may be indicated as eligible for IPD sharing they are not always obtainable upon request – inaccessible in up to 20% of instances. • IPD packages provided to researchers are often subject to excessive delays, bureaucratic burdens and substantial redactions which can impact usability.
Key Limiters to Access	• Administrative burdens that impact regulator abilities to share CSRs. • Company not having a CSR sharing policy. • Company not sharing trial CSRs until completion of follow-up on longer-term outcomes. • Company employs CSR redaction methods that impact data usability.	• Company not having an IPD sharing policy. • Company not sharing trial IPD until completion of follow-up on longer-term outcomes. • IPD embargo periods. • Company manages the IPD sharing process internally. • Company employs IPD redaction methods that impact data usability.
Steps To Increase Accessibility	• Regulators and journals could consider mandating pharmaceutical companies to make key CSRs downloadable from public repositories. • All pharmaceutical companies should have CSR sharing policies, ideally sharing via open-access models. • Pharmaceutical companies should ensure their policies make all CSRs submitted to support drug approvals are accessible (irrespective of continuing follow-up).	• All pharmaceutical companies should have IPD sharing policies. • Companies should aim to only assess if trials are in scope for IPD sharing. IPD should either be openly accessible or decisions on the legitimacy of IPD requests should be made entirely by an independent panel. • Regulators and journals could consider mandating pharmaceutical companies to make IPD from any clinical trial submitted to support drug approvals should be immediately eligible for sharing (irrespective of continuing follow-up). • Companies should ensure their trial consent procedures avoid issues with IPD sharing.

Accessing CSRs from regulatory bodies presents its own set of challenges. Operational challenges led to the EMA suspending CSR sharing from December 2018 to September 2023 [55]. While Health Canada has made CSRs prospectively accessible, it has published only a limited number of

documents predating its 2019 commitment [57, 59]. Overall, the operational challenges highlighted by the FDA [56] and EMA [55] indicate that regulators may benefit from requesting sponsors to make CSRs available at registration via sponsor-backed international platforms, like clinicaltrials.gov.

In addition to regulatory actions, proactive steps have been taken by many pharmaceutical companies to voluntarily establish CSR transparency policies. However, the absence of standardised, binding guidelines has led to inconsistent transparency commitments between companies. The most common reasons cited by sponsors for trial CSRs being ineligible to request are 1) the company does not have a CSR sharing policy, and 2) the company's data transparency policy does not enable request of CSRs for trials still in follow-up [59]. Upon registration of a drug, all trial CSRs linked to drug labels should be immediately accessible - ideally through open-access models - regardless of the trials' follow-up status. This is particularly relevant to oncology trials, where extended survival and toxicity follow-up are common. It is also notable that practices of excessively redacting aggregate-level information from shared CSRs have been observed [14, 59]. Protecting patient anonymity is paramount, but instances, where companies unnecessarily redact anonymised information to the point of inhibiting scientific endeavours should not be accepted [14].

How to achieve comprehensive IPD sharing?

Regarding IPD sharing, current regulatory frameworks lack mandates on disclosure of trial IPD. Despite this, all of the top 20 pharmaceutical companies by revenue have voluntarily implemented IPD sharing policies [43]. However, there is considerable heterogeneity in IPD sharing eligibility criteria. Standardization of policies would greatly simplify and expand IPD accessibility. The most commonly cited reasons for IPD ineligibility by sponsors include 1) the company policy does not share trial IPD until the completion of all follow-ups, 2) the IPD is under embargo, 3) the drug is not approved by both the EMA and FDA, 4) phase 1/2 trials are out of scope, and 5) the trial's consent procedures did not cover IPD sharing [43]. Moreover, some pharmaceutical companies require the submission of a research proposal before determining IPD sharing eligibility. While such requirements can ensure responsible use of data, they may inadvertently introduce barriers to transparency and access.

Most importantly, IPD from any clinical trial informing a drug approval should be immediately eligible for sharing upon drug approval, irrespective of ongoing follow-up. If the trial data are sufficiently mature to support widespread use of the drug, then they must also be sufficiently mature to be shareable with the public for assessment. Secondly, IPD sharing rates were found to be lower when the sharing process is managed internally by the company sponsoring the trial [43]. To address the research proposal limitation, clearer stipulations - such as transparent evaluation criteria and public

disclosure of accepted proposals - should be implemented. All companies should use an independent review panel to decide whether the proposed research is legitimate and scientifically valid. Preferably, the entire process of handling the data sharing request, decision-making and access would be handled by an independent, un-conflicted third-party [14, 60].

Finally, with 90% of trial participants supporting data sharing, there is a clear need to ensure that trial consenting procedures facilitate IPD sharing [26]. In alignment with National Institutes of Health (NIH) recommendations, these consents should clearly outline data sharing intentions [61], including details on the anonymisation processes to be implemented and information on who will have access to the data and why it is important. The 2017 revisions by the US Office of Human Research Protections to the Common Rule further support the implementation of such practices by interpreting non-objection on consent forms as consent for data sharing, provided privacy is safeguarded [62]. Further, akin to the case with CSRs, while de-identification of IPD is essential, redactions should not unnecessarily impede data utility. This underscores a need for a balanced approach to data privacy [18].

Moving Forward from Historical Barriers to Data Sharing

The adoption of the PhRMA/EFPIA 'Principles of Data Sharing' in 2013 marked a significant acceleration in the industry's shift towards greater data transparency [13]. Since then, many pharmaceutical companies have successfully shared CSRs and IPD for the benefit of the community and advancing research [43, 59]. This progress has shown that early concerns regarding patient privacy and financial costs are no longer the major barriers to data sharing. Currently, the main challenges to comprehensive data sharing for industry-sponsored oncology trials stem from the variability in data sharing policies across companies and the lack of clear guidelines from regulatory agencies to ensure comprehensive data sharing.

The solutions developed in the last decade to address historical concerns are evident in today’s data sharing ecosystem. Access to IPD for clinical trials typically occurs through processes that rigorously assess the qualifications of requesting researchers [14]. When access is granted, it typically occurs within secure, password-protected environments where only de-identified data is provided for analysis in accordance with comprehensive guidelines on data anonymisation. These approaches collectively ensure very stringent protection of patient privacy and are in alignment with NIH recommendations [61]. Concerning CSRs, when companies share these documents, they are usually accessible through the same portals as IPD. Additionally, when regulators release CSRs, they make significant redactions prior to public release to ensure complete anonymisation of the materials.

Regarding costs, it is evident from NIH recommendations that data sharing is recognised as an economical strategy to maximise the impact of research, rather than being a hindrance. Drawing from the NIH's model, where NIH-funded trials are now required to be shareable, it is justifiable that regulatory bodies should consider similar requirements on pharmaceutical companies for industry-sponsored trials - particularly when such trials underpin the approval of medicines commonly used in the community.

Conclusion

In conclusion, this chapter underlines the substantial advancements made in the sharing of CSRs and IPD for industry-sponsored oncology trials over the past decade. It also highlights the critical improvements that are necessary to facilitate practice-changing systematic reviews and meta-analyses that can influence and advance personalised patient care. Key steps for improving data transparency, as evidenced by current research, include standardising data sharing policies to stipulate the immediate accessibility of CSRs and IPD at the time of the drug approvals, establishing independent review processes for decision-making on IPD sharing, and striking a patient-centred balance between data security and the usability of shared information. Notably, achieving these objectives is a collective responsibility involving multiple stakeholders, including journals, regulatory bodies, professional organizations, and the pharmaceutical industry, all of whom are in positions to implement and recommend policies that promote data sharing. Additionally, it is important to recognise that while this article has primarily focused on the data sharing practices of the pharmaceutical industry, the IPD perspectives are equally applicable to non-industry/academic trial sponsors.

CHAPTER 4: THE STATE OF INDIVIDUAL PARTICIPANT DATA SHARING FOR HIGHEST REVENUE MEDICINES

Note: This chapter is adapted from a published manuscript. Modi ND, Li LX, Logan JM, Wiese MD, Abuhelwa AYA, McKinnon RA, Rowland A, Sorich MJ & Hopkins AM. The state of individual participant data sharing for the highest-revenue medicines. Clinical Trials. 2024. DOI: https://doi.org/10.1177/17407745241286147

Contributions: In this publication, I, Modi ND, contributed significantly to the writing and editing of the manuscript. Specifically, I was responsible for 90% of the writing and editing. Li LX, Logan JM, Wiese MD, Abuhelwa AYA, McKinnon RA, Rowland A, Sorich MJ & Hopkins AM contributed the remaining portions of the writing and editing in varying percentages.

Introduction

Motivated by the insights from oncology, this chapter expands the focus to assess IPD sharing for the highest revenue medicines across various therapeutic areas.

Over the past decade, there has been notable emphasis by pharmaceutical companies [13], advocacy groups [63, 64], and regulatory bodies [65] regarding the importance of facilitating secondary access to detailed IPD from clinical trials. Such access is recognised as key in facilitating independent validation of research findings, enhancing scientific collaboration, and fostering trust in the pharmaceutical industry [2, 22, 63, 66, 67]. Consequently, a majority of pharmaceutical companies now have IPD sharing policies [43, 68].

The effectiveness of current IPD sharing policies used by the majority of companies requires further evaluation. A 2018 audit revealed that only 15% of industry-sponsored clinical trials had their IPD available for sharing two years after publishing the primary outcome [11]. Additionally, only 45% of the 304 industry-sponsored clinical trials supporting FDA approval of anticancer medicines between 2011 and 2021 were eligible for IPD sharing [43]. Notably, the rate of IPD sharing dropped to 35% for the top 10 oncology medicines by revenue [43]. These findings highlight that the vital IPD that underpins the safety and efficacy of medicines used by many may not be readily accessible and that the accessibility of IPD in fields beyond oncology has not been recently evaluated.

This study aimed to assess the eligibility of independent researchers to access IPD from clinical trials that supported the FDA approval of the top 30 pharmaceutical medicines by revenue of 2021.

Methods

Sample and data

This cross-sectional study assessed the percentage of clinical trials eligible for IPD sharing, which contributed to the FDA approval of the top 30 pharmaceutical medicines by revenue in 2021. Our sample comprised strategically significant medicines - those with high revenue and, consequently, high patient usage - while also ensuring a sample size of over 300 clinical trials to achieve a 95% precision level in planned statistical analyses.

In September 2022, a list of the top 30 highest revenue-generating medicines for 2021 was compiled, using data from 'Drug Discovery & Development' and 'S&P Global Market Intelligence' [69, 70]. The FDA, EMA, and Health Canada approval status of each medicine was documented. Subsequently, the most recent FDA-authorised drug labels for these medicines were obtained from the FDA Drug Database [37, 71]. From these labels, identifiers of trials with results in the labels, as well as the dates when the trial results were integrated into the drug labels were collated. Using these identifiers, we searched ClinicalTrials.gov, the EU Clinical Trials Register, and clinical trial publications sourced via PubMed and compiled a comprehensive list of information, including NCT numbers/other identifiers, start dates, final completion dates, trial phases and details of the trial sponsoring entities.

For the trial sponsoring entities, we documented their membership status with the PhRMA and/or the EFPIA. Additionally, we searched their websites to determine if they had publicly accessible IPD sharing policies. For trial sponsors with such policies, we gathered details about the data sharing process, including whether it was an internal company process or external (e.g., via platforms like vivli.org [Vivli][15], clinicalstudydatarequest.com [CSDR][16], or yoda.yale.edu [YODA][31]), as well as contact information for IPD sharing enquiries.

Determination of IPD sharing eligibility

Beginning on 1 October 2022, we sought to confirm the IPD sharing eligibility for each trial audited in this study (i.e., whether a trial was in scope of sharing). Being eligible for sharing was defined by either the identification of a public listing of the trial as being in scope for IPD sharing on a data sharing platform website, or the receipt of a positive response to a standardised enquiry (Appendix 3) from the trial sponsor (or medicine owner if different) confirming that the trial was in scope for IPD investigations by independent researchers. Ineligibility for IPD sharing was confirmed by a negative response to the enquiry (i.e., receipt of written confirmation from the trial sponsor that IPD would not be shared with independent researchers). If a trial was indicated as not eligible for IPD sharing, details of the reason(s) for ineligibility and when the trial would become eligible were requested. If no response to the initial enquiry was received, prompts were sent every 30 days after the initial

enquiry. If no response had been received by the trial sponsor (or medicine owner if different) by 180 days from the initial enquiry, the trial was deemed to be ineligible for IPD sharing.

Analysis

All statistical analyses were performed using R 4.2.2 [42]. Chi-square tests, with statistical significance set at $P<0.05$, were used to evaluate and present differences in trial IPD sharing eligibility proportions according to trial phase, trial design, process of sharing, trial start date, time since trial completion, time since trial was listed in the drug label, and whether the IPD sharing policy includes a criterion for trial completion.

Patient and public involvement

Our investigations into clinical trial IPD transparency have been significantly guided by the contributions of our dedicated consumer advisory group whom we have been working with for the past 7 years. For this project, we extend our profound appreciation to Mark Haseloff for his indispensable insights, spanning conception, design, evaluation, and communication.

Ethics

The research undertaken was assessed as negligible risk research and was confirmed exempt from requiring Flinders University Human Research Ethics Committee review.

Results

Sample and data

The FDA drug labels for the 30 highest-revenue medicines of 2021 [69, 70] included results from a total of 316 clinical trials sponsored by the pharmaceutical industry. Each of these top 30 medicines, in terms of global revenue for 2021, were also approved by the EMA and Health Canada. The sample included 4 vaccines (involving 39 trials), 4 biologic anti-cancer medicines (involving 76 trials), 6 non-biologic anti-cancer medicines (involving 38 trials), 9 biologic medicines for non-cancer diseases (involving 102 trials), and 7 non-biologic medicines for non-cancer diseases (involving 61 trials). 8 (2%) were phase 1 trials, 44 (14%) were phase 2 trials, 261 (83%) were phase 3 trials, and 3 (1%) were phase 4 trials. 192 trials (61%) started before January 1, 2014, while 124 trials (39%) started on or after January 1, 2014. The results of 145 (46%) trials were integrated into the drug labels within the last 5 years, 95 (30%) trials were added between 5 to 10 years ago, and 76 (24%) were added more than 10 years ago. The 316 trials were sponsored by 20 different pharmaceutical companies, with 16 (80%) of these sponsors being among the top 20 for global revenue in 2021 [72]. Of these 20 companies, 17 (85%) had a publicly available IPD sharing policy, and they sponsored 311 (98%)

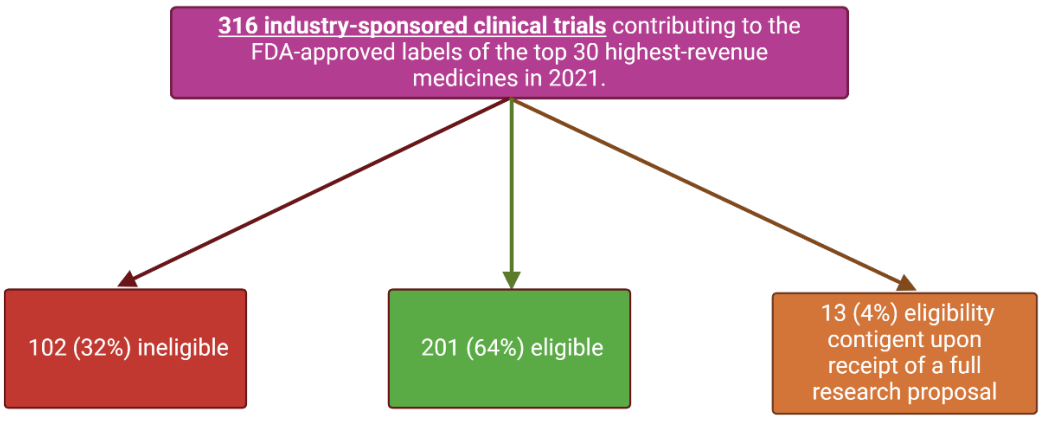
of the trials included in this study. Additionally, 14 companies had established relationships with data sharing platforms such as Vivli, CSDR, and YODA.

Of the 316 trials audited, eligibility for IPD sharing was publicly listed for 106 (34%). For the remaining 210 (66%) trials, an inquiry to the sponsor was necessary to establish their eligibility for IPD sharing, and the median response time from the initial inquiry was 42 days.

Eligibility to share

During the assessment window from 1 October 2022 to 1 April 2023, the IPD sharing eligibility of 316 clinical trials was evaluated. Out of these, 201 (64%) trials were confirmed eligible for IPD sharing, 102 (32%) were confirmed ineligible, and for 13 (4%) trials, the sponsor indicated that a full research proposal was required to assess the IPD sharing eligibility (i.e., the sponsor would neither confirm nor deny whether the trial was in scope for IPD sharing without the submission of a full research proposal) (Figure 3).

Figure 3: Diagram summarising the IPD sharing eligibility outcomes of the 316 industry-sponsored clinical trials.



Overall, among the top 30 medicines by global revenue for 2021, 19 were confirmed to have over 50% of their clinical trials, as presented in their FDA drug labels, eligible for IPD sharing (Figure 4). In contrast, Pembrolizumab, Rivaroxaban, Nivolumab and Sitagliptin/Sitagliptin-Metformin each had less than 50% of their clinical trials with results presented in their FDA drug labels eligible for IPD sharing. Furthermore, Tozinameran, Elasomeran, Apixaban and Lenalidomide had none of the trials eligible for IPD sharing.

Figure 4: Displays the percentage of IPD eligible for independent researcher request for the industry-sponsored clinical trials with results presented in the FDA-approved drug labels of 2021's top 30 highest-revenue medicines.

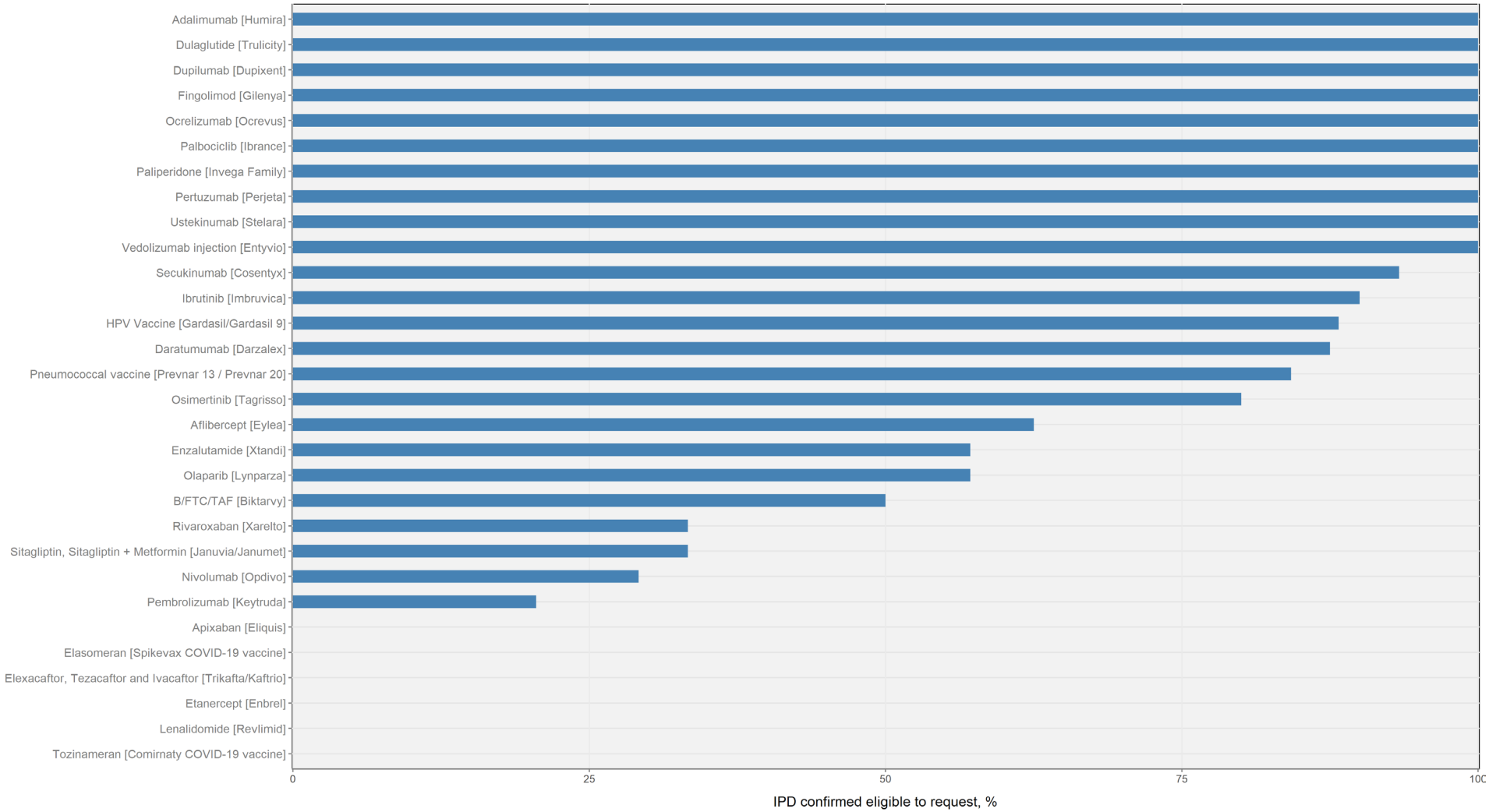


Table 6 presents the proportion of trials eligible for IPD sharing according to key descriptive subgroups. Notably, trials listed in drug labels within the last 5 years had lower eligibility rates (56%) compared to those listed 5 to 10 years ago (77%) and those listed over 10 years ago (75%) (P<0.01). Similarly, trials yet to pass the completion dates were much less likely to be eligible for IPD sharing (32%) compared to those recently completed (75% of those completed within the last 5 years, 90% of those completed within 5 to 10 years, and 70% of those completed greater than 10 years ago, P<0.01). Additionally, older trials (started before January 1, 2014) showed higher IPD sharing eligibility (76%) than those started more recently (52% for trials started after January 1, 2014, P<0.01). IPD sharing eligibility was also observed to vary with trial phase, with phase 4 trials shared at the highest rate (sharing for phase 1, 2, 3 and 4 trials was 25%, 61%, 68% and 100% respectively, P<0.01).

Table 6: Breakdown of IPD Sharing Eligibility According to Key Descriptive Subgroups

Category			Trial eligible for IPD sharing		
		Companies (n)	No	Yes	P-value*
Trial phase	Phase 1	3	6 (75%)	2 (25%)	0.03
	Phase 2	11	17 (39%)	27 (61%)	
	Phase 3	18	79 (32%)	169 (68%)	
	Phase 4	2	0 (0%)	3 (100%)	
Trial design	Non-Randomised	11	17 (40%)	26 (60%)	0.37
	Randomised	18	85 (33%)	175 (67%)	
Process of sharing	External	14	57 (25%)	171 (75%)	<0.01
	Internal	2	42 (58%)	30 (42%)	
	None Identified	2	3 (100%)	0 (0%)	
Trial start date	< 1 Jan 2014	14	44 (24%)	138 (76%)	<0.01
	≥ 1 Jan 2014	17	58 (48%)	63 (52%)	
Trial has passed final completion date	< 5 years	14	20 (25%)	61 (75%)	<0.01
	5-10 years	11	6 (10%)	54 (90%)	
	> 10 years	10	27 (30%)	63 (70%)	
	Not passed	14	49 (68%)	23 (32%)	
Time since trial listed in drug label	< 5 years	17	63 (44%)	79 (56%)	<0.01
	5-10 years	12	22 (23%)	72 (77%)	
	> 10 years	8	17 (25%)	50 (75%)	
IPD sharing policy includes a criterion for trial completion	No	8	21 (30%)	50 (70%)	0.03
	Yes	8	78 (34%)	151 (66%)	
	Has no online policy	2	3 (100%)	0 (0%)	
*Chi-sq evaluation of the distribution between trial eligible for IPD sharing "No " and "Yes"					

Beyond trial characteristics, the proportion of trials eligible for IPD sharing was also associated with characteristics of the sponsoring companies' data transparency policy (Table 6). Notably, trials sponsored by companies with no IPD sharing policies had the lowest proportion of trials eligible for IPD sharing (0%). Companies with policies that shared data via independent external platforms had a greater proportion of trials eligible for IPD sharing than compared to companies that managed their policies and processes internally (external processes 75% vs. internal processes 42%, P<0.01).

Table 7 presents a breakdown of the reasons provided by sponsors for the 115 (36%) clinical trials for which the eligibility to share IPD was not confirmed. The three most common reasons provided for confirmed ineligibility were: not in scope for sharing per policy (29 trials, 9%), the study is still ongoing (27 trials, 9%), and the clinical trial initiation or completion date predates the companies IPD sharing policy (9 trials, 3%).

Table 7: Breakdown of Clinical Trials for which Eligibility to request IPD was not confirmed

Reasons provided	Total assessed = 316
Clinical trials confirmed as ineligible for IPD request	**102 (32%)**
Not in scope for sharing per policy	29 (9%)
Study is still ongoing	27 (9%)
Clinical trial initiation or completion date pre-dates IPD sharing policy	9 (3%)
Consent form issues	8 (3%)
IPD under embargo, 18 months have not elapsed since trial completion	8 (3%)
Phase 1 trials are out of scope	7 (2%)
Results have not been published in a public registry/peer-reviewed journal	5 (2%)
Sponsor does not share IPD	3 (1%)
Co-development/contractual constraints	3 (1%)
Ongoing regulatory activities	3 (1%)
Clinical trials unable to be confirmed eligible/ineligible for IPD request	**13 (4%)**
Full research proposal required to assess eligibility to share IPD	13 (4%)
Data specified as number of trials (% of the total number of trials assessed)	

Table 8 lists the characteristics of the 30 evaluated medicines. The raw dataset generated and analysed in this study is available upon request to authors.

Table 8: Summary of the characteristics of each of the top 30 highest-revenue pharmaceutical medicines of 2021.

Drug name	Generic name	Manufacturer(s)	2021 global revenue (US$, billions)	Class	Number of clinical trials with results presented in the FDA-authorised drug label
Comirnaty COVID-19 vaccine	Tozinameran	Pfizer/BioNTech	59.1	Vaccine	2
Humira	Adalimumab	AbbVie	20.7	Biologic medicine for non-cancer disease	25
Spikevax COVID-19 vaccine	Elasomeran	Moderna	17.7	Vaccine	1
Keytruda	Pembrolizumab	Merck	17.2	Biologic anti-cancer medicine	39
Eliquis	Apixaban	Bristol Myers Squibb/Pfizer	16.7	Non-biologic medicine for non-cancer disease	8
Revlimid	Lenalidomide	Bristol Myers Squibb	12.8	Non-biologic anti-cancer medicine	7
Imbruvica	Ibrutinib	AbbVie/Johnson & Johnson	9.8	Non-biologic anti-cancer medicine	10
Eylea	Aflibercept	Regeneron/Bayer	9.2	Biologic medicine for non-cancer disease	8
Stelara	Ustekinumab	Johnson & Johnson	9.1	Biologic medicine for non-cancer disease	9
Biktarvy	B/FTC/TAF	Gilead	8.6	Non-biologic medicine for non-cancer disease	6
Xarelto	Rivaroxaban	Johnson & Johnson/Bayer	8.0	Non-biologic medicine for non-cancer disease	12
Opdivo	Nivolumab	Bristol Myers Squibb	7.5	Biologic anti-cancer medicine	24
Trulicity	Dulaglutide	Eli Lilly	6.5	Biologic medicine for non-cancer disease	10
Dupixent	Dupilumab	Sanofi/Regeneron	6.2	Biologic medicine for non-cancer disease	17
Darzalex	Daratumumab	Johnson & Johnson	6.0	Biologic anti-cancer medicine	8
Trikafta/Kaftrio	Elexacaftor, Tezacaftor and Ivacaftor	Vertex	5.7	Non-biologic medicine for non-cancer disease	2
Gardasil/Gardasil 9	HPV Vaccine	Merck	5.7	Vaccine	17
Ibrance	Palbociclib	Pfizer	5.4	Non-biologic anti-cancer medicine	2
Januvia/Janumet	Sitagliptin, Sitagliptin + Metformin	Merck	5.3	Non-biologic medicine for non-cancer disease	12
Prevnar 13 / Prevnar 20	Pneumococcal vaccine	Pfizer	5.3	Vaccine	19
Tagrisso	Osimertinib	AstraZeneca	5.0	Non-biologic anti-cancer medicine	5
Cosentyx	Secukinumab	Novartis	4.7	Biologic medicine for non-cancer disease	15
Ocrevus	Ocrelizumab	Roche	4.6	Biologic medicine for non-cancer disease	4
Enbrel	Etanercept	Amgen	4.5	Biologic medicine for non-cancer disease	10
Xtandi	Enzalutamide	Astellas	4.2	Non-biologic anti-cancer medicine	7
Invega Family	Paliperidone	Johnson & Johnson	4.0	Non-biologic medicine for non-cancer disease	17
Entyvio	Vedolizumab injection	Takeda	3.9	Biologic medicine for non-cancer disease	4
Lynparza	Olaparib	AstraZeneca/Merck	3.7	Non-biologic anti-cancer medicine	7
Perjeta	Pertuzumab	Roche	3.6	Biologic anti-cancer medicine	5
Gilenya	Fingolimod	Novartis	3.5	Non-biologic medicine for non-cancer disease	4

Discussion

This study evaluated the eligibility of independent researchers to access IPD from clinical trials supporting the FDA approval of the top 30 pharmaceutical medicines by revenue in 2021. Of the 316 trials examined, 201 (64%) trials, involving over 280,000 patients, were confirmed eligible for IPD sharing. However, 102 trials (32%), involving over 230,000 patients, were confirmed as ineligible for IPD sharing. Particularly concerning was the lack of IPD sharing eligibility for Tozinameran, Elasomeran, Apixaban, and Lenalidomide and that Sitagliptin-Sitagliptin/Metformin, Nivolumab, Rivaroxaban and Pembrolizumab had less than 50% of their trials eligible for sharing.

Recognising the importance of clinical trial data sharing, many pharmaceutical companies have committed to promoting transparency and collaboration by sharing de-identified IPD over the past decade [13]. Prior to the present study, the largest structured assessments of independent researchers' eligibility to request industry-sponsored trial IPD were conducted by Murugiah et al [44], Hopkins et al [11], and Modi et al [43]. In 2016, Murugiah et al found that IPD were eligible for sharing from ~25% of large cardiovascular trials [44]. In 2018, Hopkins et al broadened the scope beyond the cardiovascular setting, documented that ~15% of clinical trials were eligible for IPD sharing two years after publication of the primary results [11]. Positively, Modi et al's 2022 study indicated that IPD sharing eligibility was ~45% for trials that underpinned the approval of anticancer medicines in the preceding decade. Nevertheless, it was a point of concern that IPD sharing eligibility for the highest revenue-generating anticancer medicines was notably lower at 35% [43]. Significantly, the current study demonstrates that 64% of clinical trials supporting the FDA approval of the top 30 pharmaceutical medicines by revenue in 2021 are eligible for IPD requests, indicating substantial growth in the data sharing ecosystem, while also revealing room for improvement.

Our study highlights several factors that could significantly improve IPD sharing practices. Notably, companies using independent external platforms for IPD sharing had a higher proportion of eligible trials compared to those using internal processes or without any IPD sharing policies (75% vs. 42% vs. 0%, respectively; $P<0.01$). Additionally, our analysis identified a significant variance in IPD sharing eligibility based on the completion status of trials as indicated on ClinicalTrials.gov. Specifically, trials not yet marked as completed were substantially less likely to have their IPD eligible for sharing compared to those marked as completed (32% vs. 78%, respectively; $P<0.01$).

The implementation of public data sharing policies by most pharmaceutical companies over the past decade is a positive step toward transparency. However, the effectiveness of these policies hinges on their ability to ensure access to key IPD. The trials we examined are pivotal, as they contain results that contributed to the approval of the evaluated medicines by the FDA, EMA, and Health

Canada. While extended follow-up in clinical trials is essential for comprehensive safety and efficacy data, this should not be a pretext for delaying or denying access to the IPD that has already played a role in a medicine's approval [14, 43]. . Any hindrance in accessing this pivotal data undermines the very purpose of public data sharing policies designed to promote transparency. Accordingly, our study points to a necessary standardization in data sharing policies - a clear need for policies to stipulate that all IPD, especially that which underpins the results outlined in drug labels, should be immediately available for sharing upon the medicine's registration. Moreover, our findings reveal that data sharing efficacy was higher among companies that engaged external parties in their data sharing processes. This finding aligns with current discussions in the literature about mechanisms to improve data sharing practices by both industry and non-industry trial sponsors [14]. Such discussions often advocate for either the adoption of open-access IPD sharing models to minimise bureaucracy burdens or the implementation of processes managed by external parties as a means to minimise conflicts of interest and thus potentially improve IPD sharing rates [14, 65, 73].

The strength of our study lies in its focus on high-revenue medicines, offering valuable cross-sectional analysis of medicines with high patient usage that are supported by substantial funding. This approach facilitates an insightful evaluation of IPD sharing practices among major pharmaceutical companies, equipped with resources to facilitate data sharing. While our findings primarily draw from data on the top 30 medicines by revenue for 2021, the enduring high-revenue status of most of these medicines into 2022 validates the relevance of our findings [74]. However, this specific focus on high revenue medicines may have narrowed the scope of our findings, as they may not encompass the data sharing practices for all medicines or represent the practices of smaller pharmaceutical companies. Furthermore, it should be noted that for 13 trials (4%) the ability to indicate the eligibility to share IPD was contingent upon submitting a research proposal. In these instances, our study was unable to determine their current sharing eligibility. To improve this process, clearer stipulations - such as transparent evaluation criteria and public disclosure of accepted proposals - should be implemented to ensure that the requirement for research proposals does not impede open access to data. This ambiguity underscores the importance of emerging literature and recommendations from the ICMJE recognising the importance of transparent data sharing plans as part of trial registration [64].

In conclusion, this study shows that 64% (201 out of 316) of clinical trials supporting the FDA approval of the top 30 revenue-generating medicines in 2021 are eligible for IPD sharing. This finding indicates substantial progress in the pharmaceutical industry's data sharing practices over the past decade. To build on this progress, we advocate for key strategies: firstly, the adoption of either open-access IPD sharing models or the management of IPD sharing processes by independent parties;

and secondly, ensuring the immediate eligibility for sharing of primary outcome IPD critical to medicine approvals. Implementing these strategies aims to enhance the accessibility of essential data, and uphold the commitments made to clinical trial participants, who often join trials understanding that, while they may not benefit directly, their participation will help in advancing patient care. Therefore, it becomes a collective duty among all involved stakeholders to respect this commitment by maximising the potential for scientific discovery and advancement.

CHAPTER 5: A 10-YEAR UPDATE TO THE PRINCIPLES FOR CLINICAL TRIAL DATA SHARING BY PHARMACEUTICAL COMPANIES: PERSPECTIVES BASED ON A DECADE OF LITERATURE AND POLICIES

Note: This chapter is adapted from a published manuscript. Modi ND, Kichenadasse G, Hoffman TC, Haseloff M, Logan JM, Veroniki AA, Venchiarutti RL, Smit AK, Tuffaha H, Jayasekara H, Manning-Bennet A, Morton E, McKinnon RA, Rowland A, Sorich MJ & Hopkins AM. A 10-year update to the principles for clinical trial data sharing by pharmaceutical companies: perspectives based on a decade of literature and policies. BMC Med. 2023; 21, 400. DOI: https://doi.org/10.1186.s12916-023-03113-0

Contributions: In this publication, I, Modi ND, contributed significantly to the research design and the writing and editing of the manuscript. Specifically, I was responsible for 75% of the research design, and 85% of the writing and editing. Hopkins AM contributed to 10% of the research design and 5% of the writing and editing. Kichenadasse G, Hoffman TC, Haseloff M, Logan JM, Veroniki AA, Venchiarutti RL, Smit AK, Tuffaha H, Jayasekara H, Manning-Bennet A, Morton E, McKinnon RA, Rowland A and Sorich MJ contributed the remaining portions of the research design and writing and editing in varying percentages.

Introduction

Chapters 2 and 3 of this thesis highlighted significant gaps in IPD sharing for oncology medicines and explored the broader landscape and future prospects of CSRs and IPD sharing within oncology. Chapter 4 extended this assessment across various therapeutic areas, examining the top 30 revenue-generating medicines and revealing substantial advancements while identifying areas needing further improvement. Building on these insights, this chapter aims to reassess the 2013 principles endorsed by PhRMA and EFPIA by evaluating their relevance and applicability in today's context. It offers updated recommendations that address the latest challenges and innovations in drug development and regulatory frameworks. The goal is to ensure that these revised principles continue to uphold and enhance transparency, integrity, and scientific rigour within the pharmaceutical industry. Clinical trial data sharing is vital for fostering transparency, quality, scientific advancement, reducing research waste, and sustaining confidence in the pharmaceutical industry. In 2013, a large proportion of the industry, through the PhRMA and EFPIA [13], endorsed a commitment to:

(1) Share participant-level data, study-level data, and protocols from clinical trials of US and EU registered medicines with qualified researchers.

(2) Provide public access to CSRs, at a minimum synopses, from clinical trials submitted to the FDA, EMA, and EU Member States.

(3) Share summary result reports with clinical trial participants.

(4) Establish public webpages displaying the companies' data sharing policies and procedures.

(5) At a minimum, publish results from all phase 3 and any clinical trial of significant medical importance.

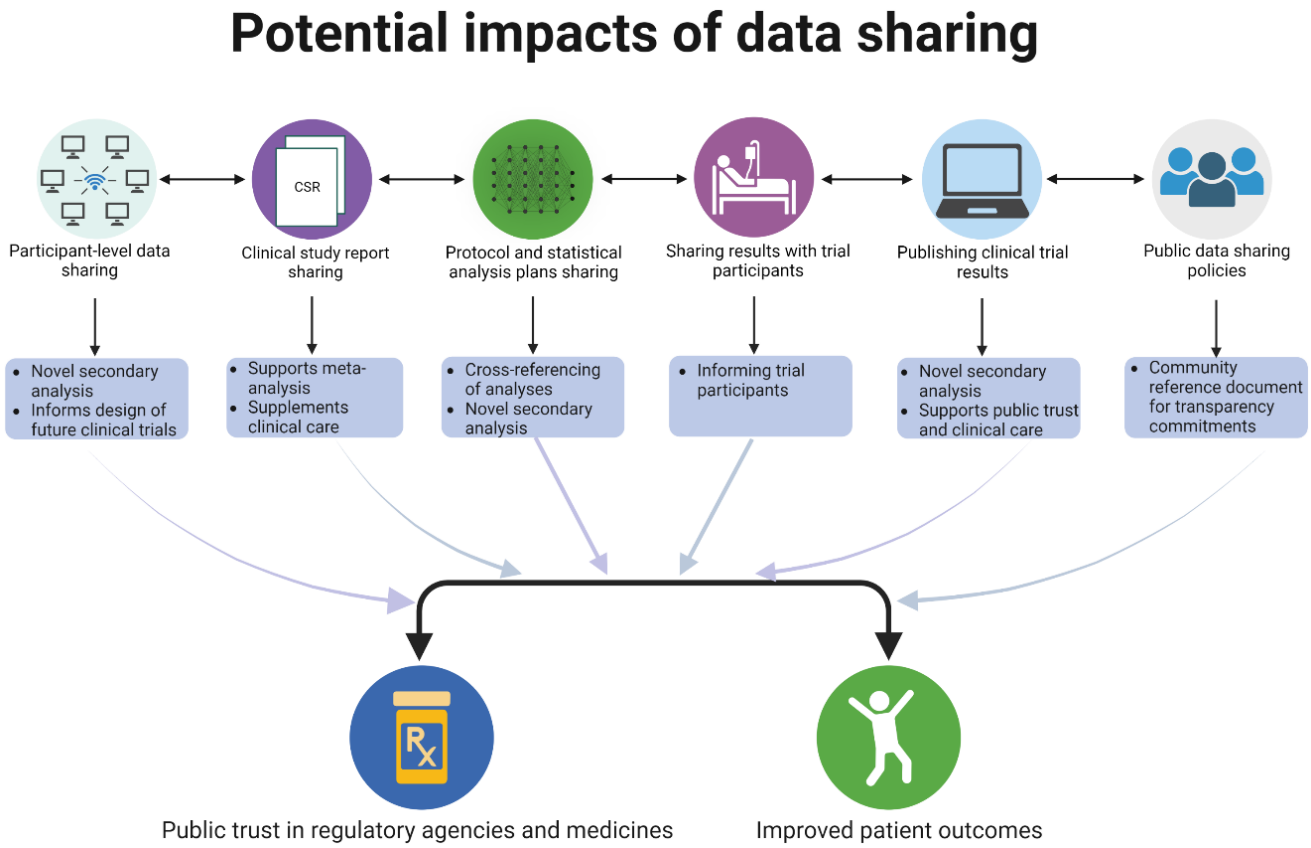
PhRMA and EFPIA members are currently at the forefront of data sharing commitments, surpassing academia, and statutory requirements. However, there is still room for further improvement and standardization of commitments to enhance communication of clinical trial results with the public, as well as to facilitate a more efficient data sharing ecosystem.

A structured literature search of Embase, PubMed, and Google Scholar was undertaken in October 2022 to identify studies that evaluated data sharing practices by pharmaceutical companies, particularly those guided by the PhRMA and EFPIA principles. The review strategy included searching terms such as 'data sharing,' 'clinical trial transparency,' 'PhRMA/EFPIA principles,' and 'clinical trial data' combined with terms specific to data sharing policies (e.g. 'individual participant data,' 'clinical study reports,' 'data registries,' etc.). The search also included relevant studies assessing the implementation of these principles across different therapeutic areas. Importantly, reference lists of key publications were examined to ensure comprehensive coverage of relevant studies.

Progress and Challenges in Clinical Trial Data Sharing

The PhRMA/EFPIA commitments marked significant progress in providing clinical trial results to participants and the general public, as well as in establishing a data sharing ecosystem that enriches the post-approval evidence base through open research conducted by independent researchers (Figure 5) [22, 24, 28, 29, 75].

Figure 5: Potential impacts of data sharing



With 18 of the current top 20 pharmaceutical companies by revenue being PhRMA/EFPIA members, the commitment holds significant weight [43]. Moreover, 15 of the top 20 companies are also TransCelerate (a collaborative network of pharmaceutical companies) members, ensuring access to guidance on collecting trial data under standardised quality conditions from the outset [76]. However, recent investigations indicate that over 50% of the clinical trials supporting the FDA approval of 115 anticancer medicines over the past 10 years were ineligible for participant-level data sharing [43]. This finding includes 90% of the clinical trials summarised in the product labels of nivolumab, pembrolizumab, and pomalidomide – this is concerning as these medicines currently rank in the top 10 anticancer medicines by global sales. Furthermore, investigations indicate that much of the participant-level data underpinning the FDA/EMA approval of COVID-19 vaccines is currently out of scope for request and will likely remain so for some time [24]. The above findings underscore an urgent need for improvements in participant-level data transparency, especially for pivotal medicines with significant medical importance.

Since 2013, policies and recommendations for sharing specific data elements have been developed by various organizations, including the FDA, EMA, Health Canada, WHO, US NIH, Institute of Medicine (now the National Academy of Medicine), White House Office of Science and Technology Policy, ICMJE, Bill and Melinda Gates Foundation, Wellcome Trust, and the GO FAIR Initiative,

among others, highlighting significant developments in the data sharing landscape [2, 49, 56, 57, 60, 77-84]. Despite these developments, the 2013 PhRMA/EFPIA principles still serve as a significant point of reference within the data sharing policy webpages of many pharmaceutical companies [13].

Enhancing Data Sharing Practices

Drawing on a decade of literature and policy developments, this article presents perspectives from a multidisciplinary team of authors, including researchers, clinicians, and consumers. The article works towards proposing evidence-based recommendations for potential updates to the pharmaceutical industry data sharing principles established in 2013. The primary aim was to review the current literature to identify and highlight feasible, urgent next steps for enhancing the data sharing ecosystem and for promoting harmonised data sharing practices among companies. The recommendations have been formulated based on the current literature and reported experiences. However, it is acknowledged that they may not address all the challenges at hand, and continued progress will still be necessary.

Table 9 presents the recommended updates, which aim to enhance existing principles, promote harmonised data sharing practices, and establish clearer guidelines regarding which data should be shared, when it should be shared, and under what conditions. The goal is to foster the data sharing ecosystem [12, 45]. Exemplifying the feasibility of the recommendations presented in Table 9, most are currently implemented in a fragmented manner across companies. While the primary focus of this manuscript is on pharmaceutical industry data sharing practices, the perspectives are also relevant to non-industry trial sponsors and investigator-initiated trials. Additionally, this study is expected to be particularly valuable for smaller pharmaceutical companies that have less established data sharing practices [45]. Outlined below are the key literature and policy developments justifying the recommendations.

Table 9: Recommendations for updating data sharing policies

Summary of 2013 PhRMA/EFPIA principles	Recommended updates to the principles.
Participant-level data sharing with researchers	
• Pharmaceutical companies commit to sharing with qualified researchers' patient-level data from clinical trials for medicines and indications approved in the US and EU. • Each company will establish a scientific review board who are not employees of the company. • Data requests will be evaluated against a description of the data being requested; hypothesis being tested; research rationale; analysis plan; publication and posting plan; qualifications and experience of the team; and a description of conflicts of interest, including potential competitive use of the data and the source of any research funding. • Companies will implement a system to provide applicable data and protocols to help facilitate the research.	1. Participant-level data from any clinical trial result submitted to support drug approvals should be eligible for sharing (irrespective of continuing follow-up). Companies should endeavour to facilitate the sharing of clinical trials not directly supporting medicine approvals within a clearly defined timeframe of primary result completion/publication. 2. Companies should aim to only assess if trials are in scope for participant-level data sharing. All decisions on the legitimacy of a data request should be evaluated by an independent scientific review panel. 3. Companies should outline the date on which their trial consent procedures were last updated and provide an example form to avoid issues with future data sharing. 4. Companies should maintain public lists of sponsored trials that are eligible/ineligible for participant-level data sharing. 5. Where possible, companies should provide full CSRs, data dictionaries, data derivation documents, protocols, SAPs, and anonymisation guides with requests to help facilitate valid secondary research.
Sharing of clinical study reports	
• To help patients and healthcare professionals understand the results of clinical trials and the evidence used to approve a new medicine (US and EU), pharmaceutical companies will make publicly available, at a minimum, the synopses of CSRs for clinical trials. • Companies will evaluate requests for full CSRs.	1. While initiatives to share result synopses are admirable, given the extent of extra clinical information and detail contained in CSRs, full CSRs from all clinical trials submitted to support medicine approvals should be publicly available for download. 2. Subsequent versions of CSRs should be made available when prepared. 3. Both the FDA and EMA have acknowledged resource difficulties in disseminating CSRs, thus it is likely companies need to engage in processes that facilitate public downloads.
Sharing of protocols and statistical analysis plans	
• Pharmaceutical companies commit to sharing with qualified researchers' protocols from clinical trials for medicines and indications approved in the US and EU.	1. Companies need to make SAPs and protocols of all published clinical trials publicly available, and consideration should be given to sharing within six months of enrolling the first participant. 2. Updated versions of SAPs and protocols should be available when prepared.
Sharing results with trial participants	
• To help inform and educate patients about the clinical trials in which they participate, pharmaceutical companies will work with regulators to adopt mechanisms for providing a factual summary of clinical trial results to research participants.	1. All trial participants should be provided a lay summary reporting trial results within 12 months of primary outcome completion. These lay summaries should also be made publicly available at that time. 2. Subsequent summaries should be prepared for follow-up outcomes. 3. Study protocols should include plans for lay summaries.
Publishing clinical trial results	
• All clinical trials should be considered for publication irrespective of whether the results were positive or negative. At a minimum, results from all phase-3 trials and any trial results of significant importance should be published.	1. All clinical trials must have result summaries published to the trials registry site within 12 months of the primary outcome completion, with efforts to make a scientific journal publication available within the same timeframe. 2. Result summaries and scientific journal publications should occur for follow-up outcomes. 3. Publishing of clinical trial results should occur regardless of study outcomes or phase. 4. Study protocols should include plans for publications.
Public data sharing policies	
• Companies following the 2013 PhRMA/EFPIA Principles for Responsible Clinical Trial Data Sharing will certify on a publicly available website that they have established policies and procedures to implement these data sharing commitments.	1. Towards harmonising terminologies and processes, companies should have public data sharing policies providing precise and detailed information on policies and procedures (including weblinks for access) to sharing participant-level data, full CSRs, protocol/SAPs, lay summaries, CSR synopses, reporting of results on clinical trial registries, and scientific journal publications. 2. Policies should be written with subheadings and numbered criteria, providing clear information on what data will be shared, when, and under what conditions for each data item. 3. To facilitate cross-referencing between documents, clinical trial registration and internal trial numbers/names should be included in all publications, product information leaflets, participant-level data, CSRs, protocols/SAPs, and lay summaries.
Abbreviations: clinical study reports (CSRs), European Federation of Pharmaceutical Industries and Associations (EFPIA), European Medicines Agency (EMA), European Union (EU), Food and Drug Administration (FDA), Pharmaceutical Research and Manufacturers of America (PhRMA), United States (US).	

Participant-level Data Sharing

Transparent sharing of participant-level data facilitates novel secondary analyses, avoids unnecessary study duplication, and informs future trial design [22, 24, 25, 28, 29, 75]. Participant-level data from clinical trials of newer medicines are vital as they are the centrepiece of safety and efficacy for these medicines [1, 24]. The EMA has indicated that they will implement future policies to promote participant-level data sharing [85], albeit, no US or EU regulations currently mandate participant-level data sharing from industry-sponsored medicine trials.

Nonetheless, most large pharmaceutical companies have processes to share participant-level data [43]. However, recent research indicates that approximately 50% of participant-level data supporting newly registered medicines are not eligible (i.e., in scope) for request [11, 24, 43, 44]. To expand data sharing, research suggests that participant-level data from any clinical trial underpinning a product label or submitted to the FDA or EMA for drug approval should be immediately eligible for sharing [24, 43]. Sharing this participant-level data should not be restricted by the clinical trial having long-term follow-up. While long-term follow-up is crucial to understanding longer-term safety and efficacy, it should not prevent sharing of result data that are responsible for the medicines approval [24, 43]. Pharmaceutical companies should also facilitate sharing of clinical trials that do not directly support medicine approvals, within a well-defined timeframe after the primary results are completed or published to reduce research waste [24, 43].

Decisions on the legitimacy of independent data requests, including the hypotheses tested, the research rationale, the analysis plan, the publication plan, and the qualifications of the research team should be made by independent scientific review panels [60]. To facilitate these review processes, it is important to establish mechanisms that provide training to independent individuals, enabling them to develop a deep understanding of the technical, legal, and scientific aspects required to assess data requests [86, 87]. The objective is to establish a pool of independent reviewers, enabling pharmaceutical companies to limit their role to simply determining the sharing eligibility of the requested participant-level data. Towards this, pharmaceutical companies should be aiming to maintain up-to-date, publicly accessible registers documenting the sharing eligibility of their clinical trials [84]. This should include a specific indication of clinical trials that are ineligible, along with clear reasons outlining why and when trials will become eligible. Among the various reasons for ineligibility, consent form issues have been identified as a major concern. To this issue, company webpages should provide clear information on updated consenting procedures, along with consent form examples [84].

Data protection and security must be a top priority for all parties, including the requestor [88]. Participant-level data sharing typically takes place on platforms requiring rigorous assessment of the requesting teams' qualifications [43, 60]. Researchers often obtain access to data in a secure, password-protected research environment from which data cannot be downloaded locally [60]. The procedures for anonymising data should align with the level of protection required. Procedures that redact key information (such as survival and adverse event data) for secondary research should be evaluated for appropriateness and necessity [60, 88]. Further, to facilitate the valid use of participant-level data, companies should enhance the findability and accessibility of clinical study reports, annotated case report forms, data dictionaries, data derivation documents, protocols, statistical analysis plans, and anonymisation guides. Such transparency, as highlighted by the FAIR (Findable, Accessible, Interoperable, and Reusable) data principles, is essential for enabling independent researchers to create detailed data requests and verify their data preparation processes when undertaking participant-level data analyses [84, 89, 90].

Independent researchers should also be committed to publishing their analyses, sharing code for reproducibility, maintaining data confidentiality, not disclosing data to unauthorised parties, and not attempting to re-identify study participants [13, 91]. Acknowledgments to data contributors and original investigators should be made in all secondary data use publications, and researchers should recognise that original investigator contributions may warrant authorship on new work [92].

Sharing of Clinical Study Reports

CSRs are standardised documents that contain detailed information (often >1000 pages) on study designs and study-level results from clinical trials, providing vastly more detail than either clinical trial result synopses or publications [93-96]. Given their comprehensive and high-quality nature, CSRs are a valuable resource for research, especially for meta- and patient-level data analyses. Furthermore, they can aid healthcare providers in making informed decisions for at-risk individuals – which can be particularly important for understanding toxicity likelihoods with newer medicines [47, 94, 97].

CSRs are often prepared as supporting documents for medicine submissions to approval and reimbursement bodies. CSR transparency has been acknowledged by the EMA, Health Canada, and the FDA as a mechanism to support public trust in regulatory processes [55-57, 97]. Both the EMA and Health Canada have regulations stating that they will publicly share CSRs submitted to them that support medicine approval decisions [55, 57]. However, resource difficulties have hindered the EMA in disseminating CSRs, and they have not been doing so since 2018 [55, 56]. Meanwhile,

the FDA has no CSR sharing policy and instead encourages sponsors to voluntarily disclose such information due to the logistic challenges it would face in implementing such a process [55, 56].

Whilst initiatives to publicly share result synopses and publications are commendable, our evaluations suggest that full CSRs from all clinical trials submitted to support medicine approvals should be publicly available for direct download, irrespective of whether the trial has continuing follow-up. Additionally, subsequent versions of CSRs should be made available as they are prepared, as new reports are often created for later data cuts. Given that there are functionalities to upload supporting documents (such as CSRs) on clinical trial registration websites [79, 98], this could be a future option for voluntary disclosure. Further, while ensuring patient anonymity is critical, companies should not endorse the practice of over-redaction in their CSR anonymisation processes [18, 99, 100].

Sharing of Protocols and Statistical Analysis Plans

Statistical analysis plans (SAPs) and protocols are essential resources for cross-referencing planned analyses and reporting of outcome/adverse event measures from clinical trials [101]. They also provide researchers with a thorough understanding of the data gathered during a clinical trial, facilitating the design of secondary data analyses [2]. The ICMJE recommends that SAPs and protocols should be reviewed when evaluating journal submissions and be made publicly available upon publication [48]. Similarly, NIH regulations (effective from 2017) indicate that SAPs and protocols should be publicly available at the time of publishing summary results [77, 78]. Notably, in 2020, both Moderna and Pfizer released detailed protocols for their COVID-19 vaccine trials, well before publishing results [102]. We propose that companies should publicly share SAPs and protocols for all published clinical trials and consider sharing them within six months of enrolling the first participant. Functionalities to upload SAP and protocol documents are available on clinical trial registries [79, 98]. Subsequent versions of SAPs and protocols should be made available when prepared (i.e., updates occur). Data management and data sharing plans should be outlined in SAPs and protocols [103].

For secondary analyses of shared data, academic institutions, and data sharing platforms should have public processes for documenting approved SAPs and requests.

Sharing Results with Trial Participants

Lay summary documents (or plain language summaries) are reports that convey clinical trial results in a simplified format for study participants and the general public [104, 105]. Sharing of such documents is recognised by regulators and companies as a mechanism to enhance public trust in

medicines [104, 106, 107]. The Declaration of Helsinki (2013) mandates that all participants 'should be given the option of being informed about the general outcome and results of the study' [108].

Companies should meet the lay summary requirements of the European Union Clinical Trials Regulation (EU CTR) 536/2014 (effective January 2022) [81, 107]. The regulation states, and we support, that all clinical trial participants should be provided a lay summary reporting the results of the clinical trial within 12 months of primary outcome completion [81, 107]. Subsequent summaries should be prepared for collected follow-up data. EU CTR indicates all lay summaries should be made publicly available. Towards best practices, preparation, and dissemination plans for lay summaries should be included in study protocols [104].

Publishing Clinical Trial Results

The Declaration of Helsinki (2013) mandates that results from human studies should be made publicly available [108]. US and EU regulations now require the publishing of clinical trial result summaries to ClinicalTrials.gov and the Clinical Trial Information System, respectively, within 12 months of primary outcome completion [77, 80, 109]. Requests have also been made to make scientific journal publications available in the same timeframe [49]. We propose that the dissemination of result publications should not depend on clinical trial outcome or phase [49] and should cover all follow-up data. Furthermore, consistency of results presentations between publications, regulatory evaluations, and product information leaflets should be ensured [110].

Public Data Sharing Policies

Pharmaceutical companies should have publicly available webpages detailing their data sharing policies, procedures, and commitments [13]. Detailed public policy information has been linked to improved clinical trial transparency [11, 43, 45, 46]. Table 8 outlines our perspectives on essential policy updates for data sharing based on emerging literature over the past decade. To implement these updates, companies should establish clear public policies for sharing participant-level data, full CSRs, protocol/SAPs, lay summaries, CSR synopses, reporting of results on clinical trial registries, and journal publications [2]. These are among the critical domains of data sharing advocated by the (now) National Academy of Medicine [2].

We recommend that data sharing policies should be written in a standardised format, including sub-headings for each data item, with numbered criteria for easy referencing by independent scientific review panels. Public registers of data sharing requests and decisions should be kept up-to-date [60]. Additionally, companies should have a register of their clinical trials that are eligible for data sharing and those that are not [84]. The register should specify the eligibility criteria and procedures

for accessing participant-level data, full CSR, protocol/SAPs, lay summary, CSR synopsis, reporting of results on clinical trial registries, and scientific journal publications for every clinical trial [2, 84].

To facilitate cross-referencing and linkage between documents, company processes should aim to include both clinical trial registration numbers and internal trial numbers/names in all publications, product information leaflets, participant-level data, CSRs, protocols/SAPs, and lay summaries [84, 111]. This cross-referencing between documents is currently undertaken poorly by most companies.

Future Directions

While the primary aim of the article was to highlight feasible, urgent next steps for enhancing the data sharing ecosystem, it is acknowledged that continued progress will still be necessary even if all the recommendations put forward are adopted. Looking ahead, the clinical trial data sharing landscape holds tremendous potential for fostering new scientific discoveries and informing decision-making [73, 112-114]. Notably, Vivli alone as a participant-level data sharing platform has facilitated the publication of over 180 research works over the past 5 years, an output that has increased from 2 manuscripts in 2019 to 85 in 2022 [115]. However, to fully realise the potential impact of the data sharing ecosystem, it will be important for all clinical trial sponsors and investigators, including non-industry trial sponsors, to take significant steps in improving standards.

It is also acknowledged that at present the data sharing landscape is fragmented in many aspects [3]. In the future, there is hope for better utilization of public clinical trial registries as valuable resources for prospectively acknowledging the sharing eligibility of participant-level data, as well as facilitating public access to CSRs, protocol/SAPs, lay summaries, result publications, annotated case report forms, data dictionaries, data derivation documents, and anonymisation guides [116-118]. At present the reporting and accessibility of these documents is somewhat disparate between companies, and the sharing eligibility of participant-level data for specific clinical trials is often not outlined prospectively.

Another consideration is the potential to centralise or transition participant-level data sharing to more open-access models. Undoubtedly, needing to access different platforms/servers (e.g. CSDR [16] and Vivli [15]) is a limiter to the effectiveness of undertaking participant-level data meta-analyses for investigations involving multiple companies. Considerations should be given to whether more open models, could facilitate crowd-sourced insights as well as minimising administrative burdens. Nonetheless, even with such a system there is still a need for mechanisms that ensure the quality of outputs.

To enhance data sharing practices, there is a need for better methods to assess and distinguish between good and bad data sharers. A valuable step towards achieving this would be the implementation of improved meta-metrics on clinical trial data sharing. Currently, the best option for comparing the transparency practices of pharmaceutical companies is 'The Good Pharma Scorecard' [68], however, it primarily ranks policies rather than comparing the outputs and performances of the companies. It is suggested that 'The Good Pharma Scorecard' could be significantly enhanced by incorporating insights into meta-metrics such as the total number of data requests received, the number of approved requests, and the number of citable public outputs facilitated for each company. This would offer a more comprehensive and transparent evaluation of data sharing efforts, enabling better recognition of companies with commendable metrics, and encouraging others to meet the standards of their competitors.

Conclusions

Data sharing plays a vital role in fostering scientific progress and supporting well-informed decisions in clinical practice. Table 9 presents policy and process updates to enhance accessibility and transparency of participant-level data, CSRs, protocol/SAPs, lay summaries, and result publications from clinical trials. Implementing these principles will require resources, time, and commitment, and we acknowledge that new issues and areas for improvement may arise [119-121]. Nonetheless, these achievable suggestions aim to facilitate the development of a data sharing ecosystem that prioritises science and patient-centred care. Meeting these commitments is in the best interest of all institutions involved in clinical trials, including companies, universities, PhRMA/EFPIA, medical societies, advocacy groups, regulators, funders, and journals, because the ultimate goal is to ensure efficient resource utilization, foster scientific advancement, and facilitate the best decisions for patients.

TRANSITION FROM POLICY EVALUATION TO PRACTICAL APPLICATIONS

Following the evaluation of current data sharing principles and identification of key areas for policy improvement in Chapters 2-5, the attention now shifts to the practical applications of enhanced data transparency. Chapters 6 and 7 explore the direct impacts of accessible IPD on clinical practice and patient outcomes. Chapter 6 focuses on the development of clinical prediction models that enable personalised predictions of adverse effects, such as diarrhoea and neutropenia, following the initiation of abemaciclib treatment. This application highlights how granular data can inform tailored interventions, thereby optimising patient care. Building on this, Chapter 7 assesses the prognostic performance of pre-treatment patient-reported outcomes in predicting prognosis and toxicity for patients receiving contemporary treatments for HER2-positive advanced breast cancer. Together, these chapters illustrate the tangible benefits of comprehensive data sharing, demonstrating how IPD can be used to translate into actionable clinical insights and optimise treatment strategies.

CHAPTER 6: PREDICTION OF SEVERE NEUTROPENIA AND DIARRHOEA IN BREAST CANCER PATIENTS TREATED WITH ABEMACICLIB

Note: This chapter is adapted from a published manuscript. Modi ND, Abuhelwa AY, Badaoui S, Shaw E, Shankaran K, McKinnon RA, Rowland A, Sorich MJ & Hopkins AM. Prediction of severe neutropenia and diarrhoea in breast cancer patients treated with abemaciclib. The Breast. 2021; 58: 57-62. DOI: https://doi.org/10.1016/j.breast.2021.04.003

Contributions: In this publication, I, Modi ND, contributed significantly to the research design, data collection and analysis, and the writing and editing of the manuscript. Specifically, I was responsible for 80% of the research design, 85% of the data collection and analysis, and 90% of the writing and editing. Hopkins AM contributed to 10% of the research design, 5% of the data collection and analysis, and 5% of the writing and editing. Abuhelwa AY, Badaoui S, Shaw E, Shankaran K, McKinnon RA, Rowland A and Sorich MJ contributed the remaining portions of the research design, data collection and analysis, and writing and editing in varying percentages.

Introduction

Hormone receptor-positive/human epidermal growth factor 2-negative (HR+/HER2-) breast cancer (BC) represents nearly two-thirds of all breast cancer diagnosis [122, 123]. Abemaciclib is a novel cyclin-dependent kinase (CDK) 4/6 reversible inhibitor that is used in the treatment of HR+/HER2- advanced BC (ABC) [124]. Current guidelines support the use of abemaciclib as a first-line therapy either in combination with a non-steroidal aromatase inhibitor (NSAI) or fulvestrant in patients with HR+/HER2- ABC [125, 126]. Safety data emerging from the MONARCH 1, 2 and 3 clinical trials have identified diarrhoea and neutropenia (characterised by low neutrophil count) as key side effects associated with abemaciclib use [127, 128]. Diarrhoea was experienced by the majority of the patients taking abemaciclib, either as a monotherapy (90%) [129], or in combination with fulvestrant (86%) [130] or NSAI (81%) [131]. Further, neutropenia was the most commonly reported severe (grade ≥ 3) adverse event in patients treated with abemaciclib, either as monotherapy (27%) [129], or in combination with fulvestrant (27%)[130] or NSAI (21%) [131].

The regulatory approval and existing literature present limited information about risk factors associated with developing diarrhoea and neutropenia in patients initiating abemaciclib [132, 133]. Development of clinical prediction models of diarrhoea and neutropenia using routinely collected clinicopathological data following abemaciclib therapy may assist clinicians in providing personalised toxicity risks. These models can also enable clinicians to understand patients needing increased monitoring or pre-emptive strategies to manage toxicities – ultimately allowing patients to remain on beneficial treatments for longer [134, 135]. The study aimed to develop clinical prediction models that allow personalised predictions of diarrhoea and neutropenia following abemaciclib initiation.

Materials and Methods

Patient Population

IPD from Eli Lilly sponsored clinical trials MONARCH 1 [NCT02102490][129], MONARCH 2 [NCT02107703][128, 130] and MONARCH 3 [NCT02246621][131, 136] was utilised in this secondary analysis study. Data were accessed according to Eli Lilly policy and has been made available through Vivli, Inc (www.vivli.org). Secondary analysis of anonymised IPD was exempted from review by the Southern Adelaide Local Health Network, Office for Research and Ethics as it was classified as minimal risk research.

MONARCH 1 is a phase 2 single-arm clinical trial including patients with HR+/HER2- ABC enrolled to 200 mg of abemaciclib twice daily [129]. MONARCH 2 is a phase 3 clinical trial including patients with HR+/HER2- ABC Randomised (1:2) to either placebo/abemaciclib (200 mg twice daily on initiation for some patients who then underwent mandatory dose reduction to 150 mg twice daily; all other patients dosed 150 mg twice daily) in combination with fulvestrant (500 mg on day 1 and 15 of cycle 1, and on day 1 of all subsequent 28-day cycles) [128, 130]. MONARCH 3 is a phase 3 clinical trial including patients with HR+/HER2- ABC Randomised (1:2) to either placebo/abemaciclib (150 mg twice daily) in combination with a nonsteroidal aromatase inhibitor (1 mg of anastrozole or 2.5 mg of letrozole once daily on every day of the 28-day cycle) [131, 136].

Predictors and Outcomes

Adverse events were reported in all trials using NCI CTCAE (National Cancer Institute Common Terminology Criteria for Adverse Events) version 4.0 [129],[128, 130],[131, 136]. Primary assessed outcomes were the development of abemaciclib induced (as reported by the study investigators) grade ≥ 3 diarrhoea and grade ≥ 3 neutropenia occurring within 365 days of therapy initiation. Assessed pre-treatment variables were selected based on availability, prior evidence, and biological plausibility. Assessed pre-treatment variables included age (years), ECOG performance status (ECOG PS [ECOG PS is a tool that evaluates the daily living abilities of patients, ranging from 0 to 5, with higher scores indicating greater disability])[137], race (Asian or Non-Asian), weight (kg), body mass index (BMI), liver metastasis, bilirubin count, alkaline phosphatase count, albumin count, white blood cell (WBC) count, neutrophil count, aspartate aminotransferase count, prior neoadjuvant/adjuvant endocrine therapy or chemotherapy, and concomitant use of antidiarrhoeals or opioids. Variables such as age and weight were dichotomised for improved model fit and clinical relevance (details provided in the statistical analysis section).

Statistical Analysis

Univariable Cox proportional hazard analysis was used to assess the association between pre-treatment variables and abemaciclib induced toxicities. Associations were reported as hazard ratios (HR) with 95% confidence intervals (CI). Statistical significance was set at a threshold of $P<0.05$ and was determined via the likelihood ratio test. Continuous variables were categorised based on model fit, observed non-linearity, prior evidence, and clinically interpretable cut-points. All analyses were stratified by treatment arm and abemaciclib dose. Prediction performances were assessed via the concordance statistic (*c*-statistic). Multivariable prediction models were developed using a stepwise forward inclusion, backwards deletion process. On forward inclusion, variables were included based on statistical significance and the greatest improvement in the *c*-statistic at each step. On backwards deletion, variables were excluded if they did not increase the *c*-statistic by 0.01. The backwards elimination process was conducted with a focus on selecting the minimal number of predictors that maintained prediction performance. To facilitate clinical use, final multivariable prediction models were converted into a toxicity risk scoring tool with the variable coefficients scaled to a point score. The tool was internally validated using machine learning. Specifically, the potential for model overfitting and robustness of variable importance were assessed using a random forest with a 10 fold cross-validation, repeated 10 times, approach [138]. Kaplan-Meier analysis was used for plotting and estimating probabilities. All data analysis was conducted using R version 3.6.2 [42].

Results

Patient population

Data were available from 900 patients. Pre-treatment patient characteristics are presented in Table 10. Median follow-up was 21 months [95% CI: 20 – 22] in MONARCH 1, 18 months [18 – 19] in MONARCH 2, and 26 months [26 – 27] in MONARCH 3.

Table 10: Summary of pre-treatment characteristics for patients who received Abemaciclib by Study

	Total No. 900	MONARCH1 No. 132	MONARCH2 No. 441	MONARCH3 No. 327
Actual treatment (arm)				
Abemaciclib-150mg + Fulvestrant-500mg	320 (36%)	0 (0%)	320 (73%)	0 (0%)
Abemaciclib-150mg + NSAI	327 (36%)	0 (0%)	0 (0%)	327 (100%)
Abemaciclib-200mg	132 (15%)	132 (100%)	0 (0%)	0 (0%)
Abemaciclib-200mg + Fulvestrant-500mg	121 (13%)	0 (0%)	121 (27%)	0 (0%)
Abemaciclib Dose				
150 mg	647 (72%)	0 (0%)	320 (73%)	327 (100%)
200 mg	253 (28%)	132 (100%)	121 (27%)	0 (0%)
Abemaciclib Combination				
Fulvestrant	441 (49%)	0 (0%)	441 (100%)	0 (0%)
Monotherapy	132 (15%)	132 (100%)	0 (0%)	0 (0%)
NSAI	327 (36%)	0 (0%)	0 (0%)	327 (100%)

Age (years)				
≤ 70	693 (77%)	108 (82%)	351 (80%)	234 (72%)
> 70	207 (23%)	24 (18%)	90 (20%)	93 (28%)
ECOG PS				
0	528 (59%)	73 (55%)	264 (60%)	191 (58%)
1+	370 (41%)	59 (45%)	175 (40%)	136 (42%)
Missing	2 (<1%)	0 (0%)	2 (<1%)	0 (0%)
Race				
Non-Asian	576 (64%)	118 (89%)	262 (59%)	196 (60%)
Asian	253 (28%)	2 (2%)	148 (34%)	103 (31%)
Missing	71 (8%)	12 (9%)	31 (7%)	28 (9%)
Weight (kg)				
≥ 60	576 (64%)	91 (69%)	270 (61%)	215 (66%)
< 60	323 (36%)	41 (31%)	171 (39%)	111 (34%)
Missing	1 (<1%)	0 (0%)	0 (0%)	1 (<1%)
Body Mass Index (kg/m^2)				
Normal/underweight	413 (46%)	64 (48%)	202 (46%)	147 (45%)
Obese	215 (24%)	31 (23%)	102 (23%)	82 (25%)
Overweight	260 (29%)	35 (27%)	133 (30%)	92 (28%)
Missing	12 (1%)	2 (2%)	4 (1%)	6 (2%)
Liver metastasis	256 (28%)	93 (70%)	116 (26%)	47 (14%)
Prior neoadjuvant or adjuvant endocrine therapy	583 (65%)	100 (76%)	339 (77%)	144 (44%)
Prior neoadjuvant or adjuvant chemotherapy	492 (55%)	101 (77%)	266 (60%)	125 (38%)
White Blood Cell Count (x 10^9/L)				
≥ 6.5	288 (32%)	41 (31%)	128 (29%)	119 (36%)
5.0 - 6.5	285 (32%)	43 (33%)	129 (29%)	113 (35%)
4.0 - 4.99	197 (22%)	29 (22%)	107 (24%)	61 (19%)
< 4.0	116 (13%)	19 (14%)	67 (15%)	30 (9%)
Missing	14 (2%)	0 (0%)	10 (2%)	4 (1%)
Neutrophil Count (x 10^9/L)				
≥ 4.5	257 (29%)	39 (30%)	115 (26%)	103 (31%)
3.50 - 4.49	220 (24%)	37 (28%)	100 (23%)	83 (25%)
2.50 - 3.49	267 (30%)	38 (29%)	129 (29%)	100 (31%)
< 2.5	140 (16%)	18 (14%)	87 (20%)	35 (11%)
Missing	16 (2%)	0 (0%)	10 (2%)	6 (2%)
Alkaline Phosphatase Count (U/L)				
< 55	128 (14%)	11 (8%)	89 (20%)	28 (9%)
55-120	564 (63%)	79 (60%)	259 (59%)	226 (69%)
≥ 120	205 (23%)	42 (32%)	93 (21%)	70 (21%)
Missing	3 (<1%)	0 (0%)	0 (0%)	3 (1%)
Aspartate Aminotransferase Count (U/L)				
< 20	257 (29%)	26 (20%)	119 (27%)	112 (34%)
≥ 20	638 (71%)	106 (80%)	321 (73%)	211 (65%)
Missing	5 (1%)	0 (0%)	1 (<1%)	4 (1%)
Albumin Count (g/L)				
≥ 40	657 (73%)	82 (62%)	302 (68%)	273 (83%)
< 40	240 (27%)	50 (38%)	139 (32%)	51 (16%)
Missing	3 (<1%)	0 (0%)	0 (0%)	3 (1%)
Bilirubin Count (µmol/L)				
< 5.0	198 (22%)	34 (26%)	96 (22%)	68 (21%)
≥ 5.0	698 (78%)	98 (74%)	345 (78%)	255 (78%)
Missing	4 (<1%)	0 (0%)	0 (0%)	4 (1%)
Opioid use at Baseline	189 (21%)	29 (22%)	106 (24%)	54 (17%)
Opioid Antidiarrhoeal use at Baseline	83 (9%)	20 (15%)	36 (8%)	27 (8%)

Data are median (IQR) or number of patients (%). ECOG PS = Eastern cooperative oncology group performance status

Of the 900 patients, 750 (82%) experienced diarrhoea from abemaciclib therapy, including 110 (12%) events of grade ≥ 3 (Table 11).

Table 11: Incidence of diarrhoea and neutropenia to abemaciclib therapy by study

	Total	MONARCH1	MONARCH2	MONARCH3
	No. 900	No. 132	No. 441	No. 327
Incidence of Diarrhoea by study				
Grade 0	150 (17%)	14 (11%)	71 (16%)	65 (20%)
Grade 1	379 (42%)	59 (45%)	179 (41%)	141 (43%)
Grade 2	261 (29%)	34 (26%)	136 (31%)	91 (28%)
Grade ≥ 3	110 (12%)	25 (19%)	55 (12%)	30 (9%)
Incidence of Neutropenia by study				
Grade 0	511 (57%)	79 (60%)	242 (55%)	190 (58%)
Grade 1	35 (4%)	1 (1%)	22 (5%)	12 (4%)
Grade 2	131 (15%)	19 (14%)	59 (13%)	53 (16%)
Grade ≥ 3	223 (24%)	33 (25%)	118 (27%)	72 (22%)

Data are median (IQR) or number of patients (%).

The median time to grade ≥ 3 diarrhoea was 21 days with 81% of grade ≥ 3 diarrhoea events occurring within the first 365 days of treatment initiation. Abemaciclib dose (200 mg vs. 150 mg) was significantly associated with increased risk of grade ≥ 3 diarrhoea (P=0.035, Table 12).

Table 12: Risk of grade ≥ 3 diarrhoea by abemaciclib dose

	N	HR	95% CI	P-value
Abemaciclib dose				0.035
150 mg	647	1		
200 mg	253	1.80	1.04 to 3.10	

CI=confidence interval, HR=hazard ratio, N=number of subjects

No significant association of grade ≥ 3 diarrhoea was identified between abemaciclib + fulvestrant versus abemaciclib + NSAI versus abemaciclib monotherapy (P=0.648, Table 13).

Table 13: Risk of grade ≥ 3 diarrhoea by treatment strategy

	N	HR	95% CI	P-value
Treatment arm				0.648
Abemaciclib monotherapy	132	1		
Abemaciclib + Fulvestrant	441	0.88	0.49 to 1.57	
Abemaciclib + NSAI	327	0.71	0.33 to 1.53	

CI=confidence interval, HR=hazard ratio, N=number of subjects

Of the 900 patients, 389 (43%) patients experienced neutropenia from abemaciclib therapy, including 223 (25%) events of grade ≥ 3 (Table 11). The median time to grade ≥ 3 neutropenia was 29 days with 90% of grade ≥ 3 events occurred within the first 365 days of abemaciclib therapy. Abemaciclib dose (200 mg versus 150 mg) was significantly associated with an increase in the risk of grade ≥ 3 neutropenia (P=0.037, Table 14).

Table 14: Risk of grade ≥ 3 neutropenia by abemaciclib dose

	N	HR	95% CI	P-value
Abemaciclib dose				0.037
150 mg	647	1		
200 mg	253	1.51	1.03 to 2.21	
CI=confidence interval, HR=hazard ratio, N=number of subjects				

No significant association of grade ≥ 3 neutropenia was identified between abemaciclib + fulvestrant versus abemaciclib + NSAI versus abemaciclib monotherapy (P=0.237, Table 15).

Table 15: Risk of grade ≥ 3 neutropenia by treatment strategy

	N	HR	95% CI	P-value
Treatment arm				0.237
Abemaciclib monotherapy	132	1		
Abemaciclib + Fulvestrant	441	1.22	0.77 to 1.94	
Abemaciclib + NSAI	327	0.96	0.55 to 1.69	
CI=confidence interval, HR=hazard ratio, N=number of subjects				

Prediction of Grade ≥ 3 Diarrhoea

On univariable analysis, advanced age (> 70 years) was significantly associated with an increased risk of abemaciclib induced grade ≥ 3 diarrhoea (HR [95%CI]: 1.72 [1.14-2.58]; P=0.009) – i.e., within the 23% of individuals greater than 70 years old, the risk of grade ≥ 3 diarrhoea was 1.72 times that of an individual aged 70 or below. No statistically significant association between grade ≥ 3 diarrhoea and ECOG PS, race, weight, body mass index, liver metastasis, bilirubin count, alkaline phosphatase count, albumin count, aspartate aminotransferase count, prior neoadjuvant/adjuvant endocrine therapy or chemotherapy, or concomitant use of antidiarrhoeals/ opioids were identified (Table 16), including on stepwise forward inclusion.

Table 16: Cox proportional hazards univariable associations between pre-treatment characteristics and risk of abemaciclib induced grade ≥ 3 diarrhoea

	N	HR	95% CI	P-value
Age (years)	900			0.009
≤ 70		1		
> 70		1.72	1.14 to 2.58	
Prior neoadjuvant or adjuvant endocrine therapy	900	1.51	0.96 to 2.37	0.075
ECOG PS	898			0.256
0		1		
1+		0.8	0.54 to 1.18	
Race	829			0.358
Non-Asian		1		
Asian		0.8	0.49 to 1.29	
Prior neoadjuvant or adjuvant chemotherapy	900	1.2	0.80 to 1.79	0.37
Opioid use at Baseline	900	0.85	0.53 to 1.38	0.515
Weight (kg)	899			0.521
≥ 60		1		
< 60		0.88	0.59 to 1.31	
Aspartate Aminotransferase Count (U/L)	895			0.594
< 20		1		
≥ 20		0.89	0.59 to 1.35	
Bilirubin Count (μmol/L)	896			0.649
< 5.0		1		
≥ 5.0		1.11	0.70 to 1.76	
Albumin Count (g/L)	897			0.726
≥ 40		1		
< 40		0.93	0.60 to 1.43	
Liver metastasis	900	1.05	0.67 to 1.63	0.837
Opioid Antidiarrhoeal use at Baseline	900	1.03	0.55 to 1.92	0.93
Body Mass Index (kg/m^2)	888			0.626
Normal/underweight		1		
Obese		1.02	0.63 to 1.66	
Overweight		1.24	0.79 to 1.92	
Alkaline Phosphatase Count (U/L)	897			0.76
55-120		1		
≥ 120		0.84	0.52 to 1.35	
< 55		0.96	0.55 to 1.66	

CI=confidence interval, HR=hazard ratio, N=number of subjects, ECOG PS = Eastern cooperative oncology group performance status

The probability of grade ≥ 3 diarrhoea within the first 365 days of abemaciclib dosed at 150 mg twice daily in individuals greater than 70 years old was 13% [95% CI; 7%-18%], compared to 9% [6%-12%] for those aged 70 or below (Table 17).

Table 17: Probability of grade ≥ 3 diarrhoea by age group

Time (days)	Abemaciclib 150mg + Fulvestrant/NSAI	
	Age ≤ 70 Median (%) [95% CI]	Age > 70 Median (%) [95% CI]
28	4 [2-6]	6 [2-10]
56	6 [3-8]	9 [4-13]
365	9 [6-12]	13 [7-18]

Table 18 outlines the probability of grade ≥ 3 diarrhoea within the first 365 days of abemaciclib dosed at 200 mg twice daily.

Table 18: Probability of grade ≥ 3 diarrhoea by age group

Abemaciclib 200mg ± Fulvestrant	
Age ≤ 70 Median (%) [95% CI]	Age > 70 Median (%) [95% CI]
12 [8-16]	21 [9-32]
14 [9-19]	21 [9-32]
16 [11-22]	37 [18-52]

Further exploratory analysis also identified advanced age as significantly associated with an increased risk of abemaciclib induced grade ≥ 2 diarrhoea (HR [95%CI]: 1.56 [1.24-1.95]; P<0.001).

Prediction of Grade ≥ 3 Neutropenia

The univariable analysis identified Asian race, weight, BMI, neutrophil count, alkaline phosphatase, albumin, aspartate aminotransferase and WBC count as significantly associated with the development of abemaciclib induced grade ≥ 3 neutropenia (P<0.05; Table 19).

Table 19: Cox proportional hazards univariable associations between pre-treatment characteristics and risk of Abemaciclib induced grade ≥ 3 neutropenia

	N	HR	95% CI	P-value
Age (years)	900			0.27
≤ 70		1		
> 70		0.83	0.59 to 1.16	
ECOG PS	898			0.088
0		1		
1+		1.26	0.97 to 1.64	
Race	829			<0.001
Non-Asian		1		
Asian		2.38	1.77 to 3.19	
Weight (kg)	899			<0.001
≥ 60		1		
< 60		1.69	1.30 to 2.20	
Body Mass Index (kg/m^2)	888			0.004
Normal/underweight		1		
Obese		0.56	0.39 to 0.80	
Overweight		0.78	0.57 to 1.06	
Liver metastasis	900	1.32	0.97 to 1.79	0.081
Prior neoadjuvant or adjuvant endocrine therapy	900	1.13	0.84 to 1.52	0.418
Prior neoadjuvant or adjuvant chemotherapy	900	1.3	0.99 to 1.72	0.063
White Blood Cell Count (x 10^9/L)	886			<0.001
≥ 6.5		1		
< 4.0		11.5	7.12 to 18.4	
4.0 - 4.99		5.17	3.23 to 8.28	
5.0 - 6.5		2.26	1.38 to 3.70	
Neutrophil Count (x 10^9/L)	884			<0.001
≥ 4.5		1		
< 2.5		9.6	5.91 to 15.6	
2.50 - 3.49		4.15	2.57 to 6.69	
3.50 - 4.49		1.73	1.00 to 3.00	
Alkaline Phosphatase Count (U/L)	897			0.018
55-120		1		
≥ 120		1.43	1.05 to 1.96	
< 55		1.53	1.07 to 2.19	
Aspartate Aminotransferase Count (U/L)	895			0.001
< 20		1		
≥ 20		1.73	1.24 to 2.40	
Albumin Count (g/L)	897			0.013
≥ 40		1		
< 40		1.44	1.08 to 1.91	
Bilirubin Count (µmol/L)	896			0.522
< 5.0		1		
≥ 5.0		1.11	0.80 to 1.53	

CI=confidence interval, HR=hazard ratio, N=number of subjects, ECOG PS = Eastern cooperative oncology group performance status

On forward inclusion, Asian race, ECOG PS, alkaline phosphatase, albumin, liver metastasis, and WBC count were identified as the statistically significant predictors within a full multivariable model. The backwards elimination process resulted in a final clinical prediction model for grade ≥ 3 neutropenia optimally defined by race, ECOG PS and WBC count (< 4.0 vs. 4.0-4.99 vs. 5.0-6.5 vs. ≥ 6.5 x10^9/L) (Table 20). The discrimination performance (*c*-statistic) of the final multivariable model was 0.75 (Table 20). A risk scoring tool based on the final multivariable model was developed.

Table 20: Final multivariable model of grade ≥ 3 neutropenia following abemaciclib initiation

	HR	95% CI	P-value
ECOG PS			<0.001
0	1		
1+	1.64	1.23 to 2.18	
Race			<0.001
Non-Asian	1		
Asian	2.19	1.60 to 2.99	
White Blood Cell Count (x 10^9/L)			<0.001
≥ 6.5	1		
5.0 - 6.5	2.16	1.30 to 3.59	
4.0 - 4.99	4.42	2.72 to 7.17	
< 4.0	9.90	6.07 to 16.2	

CI=confidence interval, HR=hazard ratio, ECOG PS = Eastern cooperative oncology group performance status

Clinical prediction tool for Grade ≥ 3 Neutropenia

The scores for the prediction tool were derived by scaling variable coefficients from the final multivariable model to a point score. Asian race equated to 1 risk point, ECOG PS of 1+ equated to 1 risk point, WBC count (x10^9/L) of 5.0-6.49 equated to 1 risk point, WBC count 4.0-4.99 to 2 risk points and WBC count < 4.0 to 3 risk points (Figure 6 and Figure 7).

Figure 6: Clinical prediction model of developing grade ≥ 3 neutropenia for Abemaciclib 150 mg + Fulvestrant/NSAI therapy at 56 and 365 days

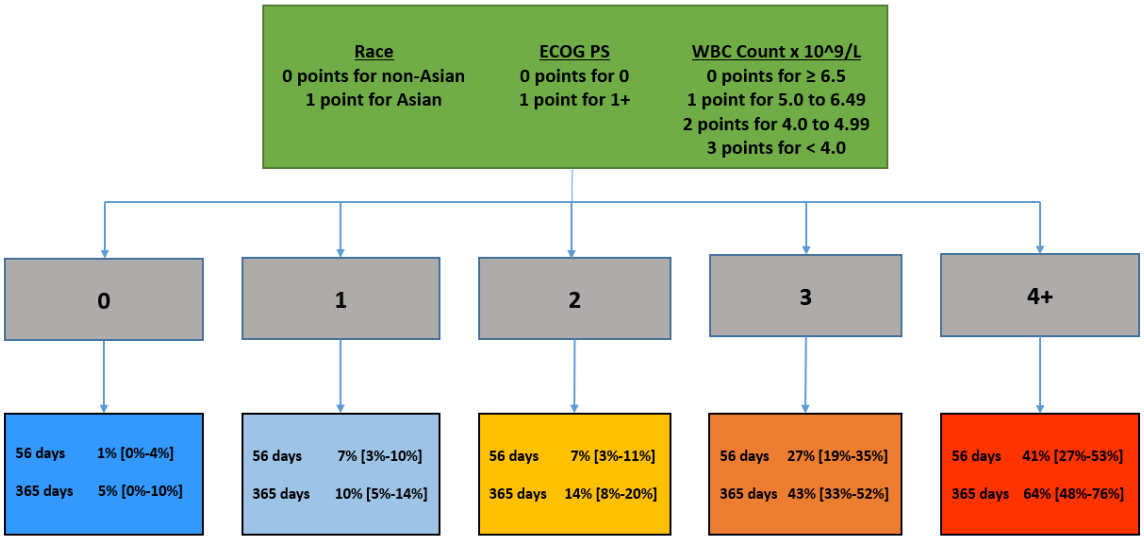
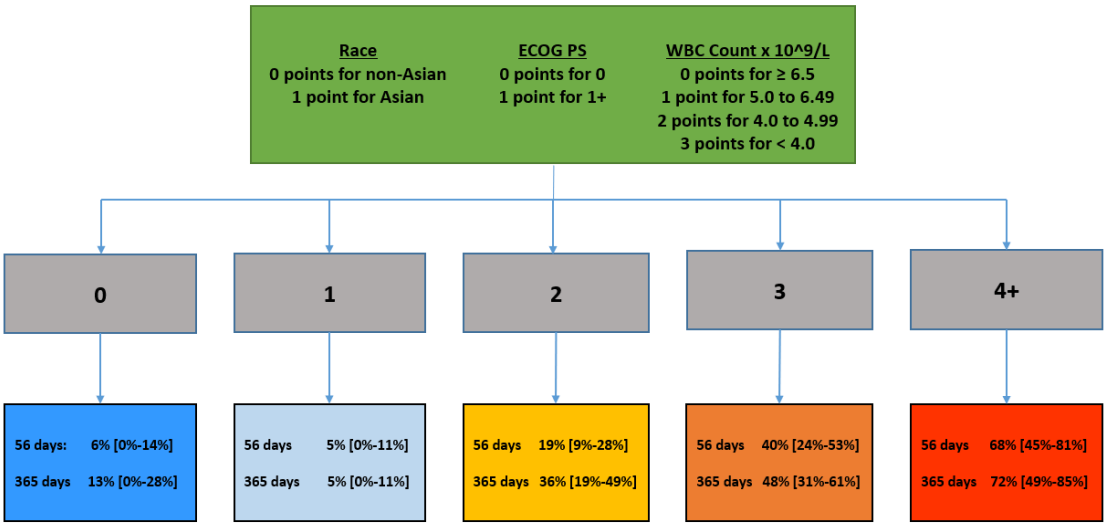


Figure 7: Clinical prediction model of developing grade ≥ 3 neutropenia for Abemaciclib 200 mg ± Fulvestrant therapy at 56 and 365 days



Patients were categorised into five subgroups according to their overall risk score (i.e., 0, 1, 2, 3, 4+). The risk scoring tool resulted in a *c*-statistic of 0.74 (Table 21).

Table 21: Cox proportional hazards univariable analysis of model based risk score and development of grade ≥ 3 neutropenia

	N	HR	95% CI	P-value
Risk Score				<0.001
0	99	1		
1	240	1.6	0.65 to 3.96	
2	219	3.49	1.49 to 8.18	
3	169	9.63	4.19 to 22.1	
4+	86	16.9	7.27 to 39.5	

CI=confidence interval, HR=hazard ratio, N=number of subjects

Table 22 and Figure 6 present the risk score tools ability to calculate probabilities of grade ≥ 3 neutropenia within the first 365 days of abemaciclib (150 mg twice daily) + fulvestrant/NSAI. Of the 11% of individuals in the highest risk subgroup (i.e., risk score = 4+) the probability of developing grade ≥ 3 neutropenia within the first 365 days of abemaciclib (150 mg twice daily) + fulvestrant/NSAI therapy was 64% [48%-76%]. Comparatively, of the 12% of individuals in the lowest risk subgroup (i.e., risk score = 0) the probability of developing grade ≥ 3 neutropenia within the first 365 days of abemaciclib (150 mg twice daily) + fulvestrant/NSAI therapy was 5% [0%-10%]. Supplementary

Table 22: Scoring metric for grade ≥ 3 neutropenia following Abemaciclib 150 mg + Fulvestrant/NSAI therapy initiation at 12 months

Neutropenia Risk Factors	Points
Asian Race	1
ECOG Performance Score 1+	1
White Blood Cell Count [5.0 to 6.49 x 10^9/L]	1
White Blood Cell Count [4.0 to 4.99 x 10^9/L]	2
White Blood Cell Count [< 4.0 x 10^9/L]	3
Maximum Risk Score	4+

Abemaciclib 150 mg + Fulvestrant/NSAI therapy	
Risk Score	Predicted Neutropenia Incidence at 12 months
0	5%
1	10%
2	14%
3	43%
4+	64%

Table 23 and Figure 7 present the risk score tools ability to calculate probabilities of grade ≥ 3 neutropenia within the first 365 days of abemaciclib (200 mg twice daily) ± fulvestrant according to defined risk groups.

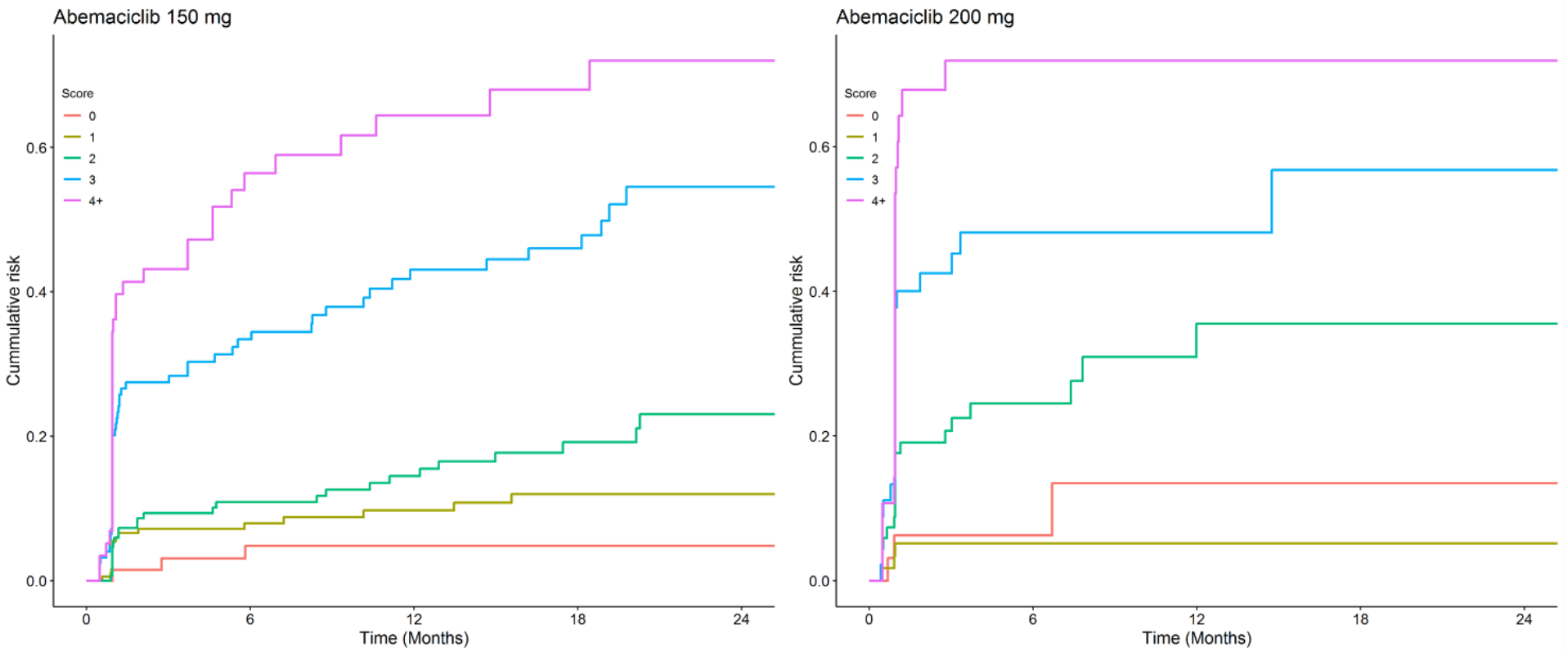
Table 23: Scoring metric for grade ≥ 3 neutropenia following Abemaciclib 200 mg ± Fulvestrant therapy initiation at 12 months

Neutropenia Risk Factors	Points
Asian Race	1
ECOG Performance Score 1+	1
White Blood Cell Count [5.0 to 6.49 x 10^9/L]	1
White Blood Cell Count [4.0 to 4.99 x 10^9/L]	2
White Blood Cell Count [< 4.0 x 10^9/L]	3
Maximum Risk Score	4+

Abemaciclib 200 mg ± Fulvestrant therapy	
Risk Score	Predicted Neutropenia Incidence at 12 months
0	13%
1	5%
2	36%
3	48%
4+	72%

The random forest approach identified race, ECOG PS, neutrophil and WBC count as the most influential variables in predicting abemaciclib induced neutropenia; confirming the validity of the variables included in the prediction tool. The discrimination performance of the repeated cross-validated random forest model was 0.75 – indicating no problems with overfitting. Figure 8 presents Kaplan-Meier plots for grade ≥ 3 neutropenia according to the predicted risk scores by assessed abemaciclib dosing strategies.

Figure 8: Kaplan Meier plots of cumulative risk of grade ≥ 3 neutropenia by Abemaciclib dose and pre-treatment risk score



Discussion

This study used large, pooled data to develop and present the first clinical prediction tool of abemaciclib induced grade ≥ 3 neutropenia in patients with HR+/HER2- ABC. The tool defined the risk of grade ≥ 3 neutropenia within the first 365 days of abemaciclib (150 mg twice daily) + fulvestrant/NSAI, which ranged from 5% to 64% according to patient race (Asian vs. non-Asian), ECOG PS (1+ vs. 0) and pre-treatment WBC count (< 4.0 vs. 4.0-4.99 vs. 5.0-6.5 vs. ≥ 6.5 $x10^9$/L). The study also identified that advanced age (70 years) was associated with an increased risk of abemaciclib induced grade ≥ 3 diarrhoea.

Neutropenia is a common side effect associated with CDK 4/6 inhibitors due to their effects on the hematopoietic bone marrow. Whilst abemaciclib has a lower incidence of neutropenia when compared to other CDK 4/6 inhibitors, neutropenia was the most commonly reported severe (grade ≥ 3) side effect associated with its use [139]. Abemaciclib induced grade ≥ 3 neutropenia is commonly managed by drug suspension and dose reduction [139]. Therefore it is important to identify the cohort of patients at high risk of grade ≥ 3 neutropenia at baseline as it can progress to neutropenic sepsis [140]. Final multivariable analysis identified race, ECOG PS and pre-treatment WBC count as the most significant predictors associated with the development of abemaciclib induced grade ≥ 3 neutropenia. The findings of the final multivariable analysis are consistent with literature identifying race [141, 142], ECOG PS [143, 144] and pre-treatment WBC count [145] as prognostic factors associated with the development of neutropenia from anticancer therapies more generally. Whilst the final risk tool had a small decline in the discriminative performance (c = 0.74) compared to the final multivariable model (c = 0.75), clinical simplicity and user-friendliness was optimised.

Prior research indicates no statistical difference in abemaciclib pharmacokinetics according to race [146], suggesting the higher risk of developing abemaciclib induced grade ≥ 3 neutropenia in Asians is likely pharmacodynamically driven. Findings from a meta-analysis on other CDK4/6 inhibitors identified no differences in neutropenia and diarrhoea risk by ethnicity [147]. Addition of ECOG PS alongside race and white blood cell count provided synergistic enhancement of model discrimination – despite ECOG PS not being a significant variable on univariable analysis.

Future research should aim to validate the presented neutropenia prediction tool for other CDK 4/6 inhibitors. Nonetheless, the presented tool has significant potential to guide clinicians in identifying patients at an increased risk of abemaciclib induced neutropenia. For example, 21% of participants were identified to have a risk score of 3+, in which the risk of grade ≥ 3 neutropenia was >40% within the first 365 days of abemaciclib (150 mg twice daily) + fulvestrant/NSAI therapy. Identifying these

patients at a substantially increased risk of neutropenia enables clinicians to consider pre-emptive strategies (e.g. prophylactic granulocyte colony stimulating factors, abemaciclib dose reductions or more stringent monitoring of white blood cell counts) to facilitate effective and safe long term abemaciclib treatment without necessitating persistent clinician-initiated interventions in the form of abemaciclib withdrawal. Minimization of persistent clinician-initiated interventions for the management side effects can also contribute to lower levels of patient anxiety to treatment [148].

Diarrhoea is a common side effect with many anticancer drugs (including with CDK 4/6 inhibitors) [149]. Abemaciclib use is associated with a higher rate of grade ≥ 3 diarrhoea compared to other CDK 4/6 inhibitors [150]. Advanced age (>70 years) was identified as the only variable associated with an increased risk of grade ≥ 3 diarrhoea, consistent with prior literature indicating that the advanced age population is at higher risk of diarrhoea from active oncological treatment [151]. The absolute difference in risk of developing abemaciclib induced grade ≥ 3 diarrhoea between the advanced and young ages was small (13% vs. 9% in the first 365 days, respectively), however, in relative terms the study was able to highlight that advanced age individuals were at 1.72 times greater risk of abemaciclib induced grade ≥ 3 diarrhoea. It is hypothesised that polypharmacy, pharmacokinetics, and pharmacodynamics changes in the advanced age subgroup, may contribute to the increased risk of abemaciclib induced grade ≥ 3 diarrhoea [152-154]. Future research should aim to elucidate the relationship between age and the risk of diarrhoea from other CDK 4/6 inhibitors and if the association is further established a stricter adherence to standardised management of diarrhoea in the form of antidiarrhoeal medications, dose reduction and drug suspension should be followed.

RCTs are the backbone of evidence-based medicine, however, strict inclusion criteria within RCTs can limit the generalizability of results [155]. Contrasting this, RCTs provide rigorous, high quality collection of adverse event data, allowing for the development of well-defined prediction tools [156]. Transparently shared IPD from the MONARCH 1, 2, and 3 trials was pooled to increase study power and enabled the development of a highly discriminatory clinical prediction tool (c = 0.74). This highlights the value of IPD sharing in creating predictive models. Effective communication of personalised and well-validated predictions of an individual's expected adverse outcomes can improve shared decision making, empower patients, and enable patients and clinicians to make better decisions regarding strategies to mitigate adverse outcomes [157]. Nevertheless, with advances in large electronic health record platforms, future opportunities to externally validate the presented tool within observational datasets of patients using abemaciclib in routine clinical care should occur – in the future this may also include evaluating the tools appropriateness for abemaciclib's use as a neo-adjuvant treatment [158].

In conclusion, the study identified advanced age as being significantly associated with an increased risk of abemaciclib induced grade ≥ 3 diarrhoea. The study also developed a clinical prediction tool based upon race, ECOG PS and WBC count for predicting abemaciclib induced grade ≥ 3 neutropenia. The developed tool offered large and substantial discrimination between subgroups, exemplifying the ability of the developed tool to inform on clinically significant difference in neutropenia risk to clinicians and patients considering abemaciclib use.

CHAPTER 7: PATIENT-REPORTED OUTCOMES PREDICT SURVIVAL AND ADVERSE EVENTS FOLLOWING ANTICANCER TREATMENT INITIATION IN ADVANCED HER2-POSITIVE BREAST CANCER

Note: This chapter is adapted from a published manuscript. Modi ND, Danell NO, Perry RNA, Abuhelwa AY, Rathod A, Badaoui S, McKinnon RA, Haseloff M, Shahnam A, Swain SM, Welslau M, Sorich MJ & Hopkins AM. Patient-reported outcomes predict survival and adverse events following anticancer treatment initiation in advanced HER2-positive breast cancer. ESMO Open. 2022; 7: 100475. DOI: https://doi.org/10.1016/j.esmoop.2022.100475

Contributions: In this publication, I, Modi ND, contributed significantly to the research design, data collection and analysis, and the writing and editing of the manuscript. Specifically, I was responsible for 75% of the research design, 80% of the data collection and analysis, and 85% of the writing and editing. Hopkins AM contributed to 10% of the research design, 10% of the data collection and analysis, and 5% of the writing and editing. Danell NO, Perry RNA, Abuhelwa AY, Rathod A, Badaoui S, McKinnon RA, Haseloff M, Shahnam A, Swain SM, Welslau M and Sorich MJ contributed the remaining portions of the research design, data collection and analysis, and writing and editing in varying percentages.

Introduction

Human Epidermal Growth Factor Receptor 2 (HER2) positive BC is an aggressive subtype of BC [159]. Evidence outlines that the emergence of targeted therapies such as trastuzumab, pertuzumab, and ado-trastuzumab emtansine (T-DM1) have improved survival outcomes in HER2-positive ABC [160-163]. Despite this, a persistent burden of unpredictable poor response remains for many patients, while others may experience significant toxicities [160, 164, 165]. Thus, predicting patients who are likely to achieve better or worse outcomes to contemporary anticancer treatment in HER2-positive ABC remains of significant interest to support shared-decision making and precision medicine.

Shared decision-making is the process in which the clinician and the patient collate and discuss the available evidence on the benefits and harms of treatments to make the most appropriate informed health decisions for the patient [166]. Shared decision-making is an essential component of providing patient-centred care [137]. ECOG PS is a clinician-interpreted tool used to evaluate the daily living abilities of patients [137]. ECOG PS is often used for oncology trial stratification and in clinical practice to evaluate prognosis and toxicity to anticancer treatment, thus supporting shared decision-making. Patient-reported outcomes (PROs) are structured self-reported tools that provide the patients' perspective and voice to their physical, social, emotional, and functional abilities [167-169]. PROs are frequently used in oncology trials as measures to evaluate treatment impacts on quality of life [170-172]. However, PRO tools are minimally used for oncology trial stratification, or in clinical practice to estimate likely benefits and harms from anticancer treatment.

PROs have shown to be of prognostic importance other cancer types (including bladder cancer, non-small cell lung cancer, melanoma) [173-180], with some studies demonstrating patient-reported physical function/well-being as more prognostic than ECOG PS [174, 177, 178, 180]. Additionally, PROs have shown the potential to detect serious adverse events earlier than clinician reporting [181]. However, the prognostic value of PROs in HER2-positive ABC has been minimally explored.

The present study aimed to evaluate the prognostic performance of pre-treatment PROs for prognosis and toxicity in patients initiating contemporary anticancer treatment for HER2-positive ABC.

Materials and Methods

Patient population

IPD from the Roche sponsored phase III clinical trials CLEOPATRA [NCT00567190, data cut: February 2014] [162, 182, 183], EMILIA [NCT00829166, data cut: December 2014] [163, 184], and MARIANNE [NCT01120184, data cut: May 2016] [185, 186] was utilised in this *post hoc* study. Data were accessed according to Roche policy and has been made available through Vivli, Inc (www.vivli.org). Secondary analysis of anonymised IPD was exempted from review by the Southern Adelaide Local Health Network, Office for Research and Ethics as it was classified as minimal risk research. CLEOPATRA included patients with HER2-positive, locally recurrent, unresectable, or metastatic BC that were treatment naïve (excluding prior hormonal therapy) in the advanced setting. Patients were randomly assigned 1:1 to receive either placebo + trastuzumab + docetaxel, or pertuzumab + trastuzumab + docetaxel [162, 182, 183]. EMILIA included heavily pre-treated patients with HER2-positive, unresectable, locally advanced, or metastatic BC with documented disease progression to trastuzumab and a taxane. Patients were randomly assigned 1:1 to either lapatinib + capecitabine or T-DM1 [163, 184]. MARIANNE included patients with HER2-positive, unresectable, progressive, or recurrent locally advanced, or metastatic BC that were treatment naïve in the advanced setting. Patients were randomly assigned 1:1:1 to trastuzumab + a taxane, T-DM1 + placebo, or T-DM1 + pertuzumab [185, 186].

Predictors and Outcomes

Pre-treatment PROs were recorded using the Functional Assessment of Cancer Therapy – Breast (FACT-B) version 4.0 questionnaire in all three studies [187]. FACT-B is a self-reported 37-item questionnaire that measures multidimensional health-related quality of life in patients with breast cancer. Responses to each question are captured on a five-point scale ranging from 0, “Not at all” to 4, “Very Much”. Answers to the 37 questions are then used to calculate subscale scores. FACT-B has five subscales: physical well-being [Score Range (0-28)], social well-being [(0-28)], functional

well-being [(0-28)], emotional well-being [(0-24)], and the breast cancer subscale [(0-40)]. Trial outcome index score [(0-96)] is a composite index of physical well-being, functional well-being, and the breast cancer subscale. The five defined subscales are also used to generate a total Fact-B score [(0-148)]. For all subscale scores, higher scores represent the patient's perception of "better" health-related quality of life. The primary evaluated predictors in this study were pre-treatment physical well-being, social well-being, functional well-being, emotional well-being, and the breast cancer subscale scores.

The primary assessed outcome was overall survival (OS), with progression-free survival (PFS) and grade ≥ 3 adverse events assessed as secondary outcomes. OS was defined as the time from randomization to the last follow-up or death from any cause - consistent across all studies. PFS was defined as the time from randomization to disease progression or death from any cause, with progression assessed using the Response Evaluation Criteria in Solid Tumours (RECIST) version 1.0 (CLEOPATRA and EMILIA) or RECIST version 1.1 (MARIANNE) [163, 183, 185]. Adverse events were reported in CLEOPATRA and EMILIA using NCI CTCAE version 3.0, and MARIANNE used NCI CTCAE version 4.0 [163, 183, 185].

Statistical Analysis

Cox proportional hazard analysis was used to assess the association between pre-treatment PROs with OS, PFS, and grade ≥ 3 adverse events. All analyses were stratified by study and treatment arm. Associations were reported as hazard ratios (HR) with 95% confidence intervals (95% CI). Statistical significance was set at a threshold of $P<0.05$ and was determined via the likelihood ratio test. Complete case analyses were conducted. Discrimination performance was assessed using the *c*-statistic. Akaike information criterion and visual checks were used to assess potential non-linear effects of continuous variables and cut-point appropriateness.

Univariable and analyses adjusted for race, sex, age, ECOG PS, BMI, estrogen/progesterone receptor status, time since the initial diagnosis, presence of visceral disease, count of tumour disease sites, prior trastuzumab/anthracycline/taxane all settings, lactate dehydrogenase concentration, and comorbidity count were conducted. Kaplan-Meier analysis was used to visually present the associations between PROs and survival/toxicity outcomes. For plotting, PROs were classified as "Poor", "Intermediate' and "Good" based on the interquartile range (IQR) of each subscale in the study population. Forest plots were used to visualise the heterogeneity in the association between PROs and survival/toxicity outcomes according to study and treatment arms.

Exploratory analysis of the prognostic performance of PROs compared to ECOG PS was conducted and assessed via the *c*-statistic. All analyses were performed using R version 3.6.2.

Results

Patient population

Data were available from 2,894 patients (Table 24) treated with contemporary therapies from CLEOPATRA, EMILIA, and MARIANNE.

Table 24: Summary of patient characteristics by study

	Total No. 2,894	CLEOPATRA No. 808	EMILIA No. 991	MARIANNE No. 1,095
The actual treatment given				
Lapatinib + Capecitabine	488 (17%)	0 (0%)	488 (49%)	0 (0%)
Pertuzumab + Trastuzumab + Docetaxel	408 (14%)	408 (50%)	0 (0%)	0 (0%)
Pertuzumab + Trastuzumab emtansine	366 (13%)	0 (0%)	0 (0%)	366 (33%)
Placebo + Trastuzumab + Docetaxel	396 (14%)	396 (49%)	0 (0%)	0 (0%)
Placebo + Trastuzumab emtansine	361 (12%)	0 (0%)	0 (0%)	361 (33%)
Trastuzumab + Docetaxel/Paclitaxel	353 (12%)	0 (0%)	0 (0%)	353 (32%)
Trastuzumab emtansine	490 (17%)	0 (0%)	490 (49%)	0 (0%)
Missing	32 (1%)	4 (<1%)	13 (1%)	15 (1%)
Arm of the clinical study				
Lapatinib + Capecitabine	496 (17%)	0 (0%)	496 (50%)	0 (0%)
Pertuzumab + Trastuzumab + Docetaxel	402 (14%)	402 (50%)	0 (0%)	0 (0%)
Pertuzumab + Trastuzumab emtansine	363 (13%)	0 (0%)	0 (0%)	363 (33%)
Placebo + Trastuzumab + Docetaxel	406 (14%)	406 (50%)	0 (0%)	0 (0%)
Placebo + Trastuzumab emtansine	367 (13%)	0 (0%)	0 (0%)	367 (34%)
Trastuzumab + Docetaxel/Paclitaxel	365 (13%)	0 (0%)	0 (0%)	365 (33%)
Trastuzumab emtansine	495 (17%)	0 (0%)	495 (50%)	0 (0%)
Treatment ARM contains Pertuzumab	765 (26%)	402 (50%)	0 (0%)	363 (33%)
Treatment ARM contains Lapatinib	496 (17%)	0 (0%)	496 (50%)	0 (0%)
Treatment ARM contains Capecitabine	496 (17%)	0 (0%)	496 (50%)	0 (0%)
Treatment ARM contains Trastuzumab	1,173 (41%)	808 (100%)	0 (0%)	365 (33%)
Treatment ARM contains TDM1	1,225 (42%)	0 (0%)	495 (50%)	730 (67%)
Treatment ARM contains Docetaxel/Paclitaxel	1,173 (41%)	808 (100%)	0 (0%)	365 (33%)
Has overall survival follow-up	2894 (100%)	808 (100%)	991 (100%)	1095 (100%)
Has progression-free survival follow-up	2894 (100%)	808 (100%)	991 (100%)	1095 (100%)
Has adverse events follow-up	2848 (98%)	777 (96%)	991 (100%)	1080 (99%)
Missing	46 (2%)	31 (4%)	0 (0%)	15 (1%)
Sex				
Male	14 (<1%)	2 (<1%)	5 (1%)	7 (1%)
Female	2,880 (100%)	806 (100%)	986 (99%)	1,088 (99%)
Age (years)				
Median (IQR)	53 (45 - 61)	54 (46 - 61)	53 (45 - 60)	53 (44 - 61)
Race				
White	1,916 (66%)	480 (59%)	732 (74%)	704 (64%)
Asian	697 (24%)	261 (32%)	180 (18%)	256 (23%)
Black or African American	124 (4%)	30 (4%)	50 (5%)	44 (4%)
American Indian or Alaska Native	20 (1%)	7 (1%)	13 (1%)	0 (0%)
Native Hawaiian or Other Pacific Islander	4 (<1%)	0 (0%)	4 (<1%)	0 (0%)
Other	131 (5%)	29 (4%)	12 (1%)	90 (8%)
Multiple	1 (<1%)	0 (0%)	0 (0%)	1 (<1%)
Missing	1 (<1%)	1 (<1%)	0 (0%)	0 (0%)
Race				
Non-Asian	2,196 (76%)	546 (68%)	811 (82%)	839 (77%)
Asian	697 (24%)	261 (32%)	180 (18%)	256 (23%)
Missing	1 (<1%)	1 (<1%)	0 (0%)	0 (0%)

ECOG PS				
0	1,852 (64%)	522 (65%)	612 (62%)	718 (66%)
1	1,027 (35%)	282 (35%)	370 (37%)	375 (34%)
2	5 (<1%)	3 (<1%)	0 (0%)	2 (<1%)
3	1 (<1%)	1 (<1%)	0 (0%)	0 (0%)
Missing	9 (<1%)	0 (0%)	9 (1%)	0 (0%)
ECOG PS				
0	1,852 (64%)	522 (65%)	612 (62%)	718 (66%)
1+	1,033 (36%)	286 (35%)	370 (37%)	377 (34%)
Missing	9 (<1%)	0 (0%)	9 (1%)	0 (0%)
Weight (kg)				
Median (IQR)	66 (57 - 77)	65 (56 - 74)	67 (59 - 78)	66 (57 - 77)
Missing	10 (<1%)	0 (0%)	9 (1%)	1 (<1%)
Body Mass Index				
Median (IQR)	26 (23 - 30)	25 (22 - 30)	26 (23 - 30)	26 (23 - 29)
Missing	25 (1%)	1 (<1%)	17 (2%)	7 (1%)
Estrogen Receptor Status				
Positive	1,434 (50%)	361 (45%)	507 (51%)	566 (52%)
Negative	1,413 (49%)	436 (54%)	467 (47%)	510 (47%)
Missing	47 (2%)	11 (1%)	17 (2%)	19 (2%)
Progesterone Receptor Status				
Positive	1,041 (36%)	251 (31%)	369 (37%)	421 (38%)
Negative	1,765 (61%)	538 (67%)	592 (60%)	635 (58%)
Missing	88 (3%)	19 (2%)	30 (3%)	39 (4%)
Time since initial diagnosis (Days)				
Median (IQR)	834 (100-1702)	587 (47-1421)	1202 (580-2200)	659 (46-1394)
Missing	60 (2%)	51 (6%)	5 (1%)	4 (<1%)
Visceral disease site at baseline	2,050 (71%)	630 (78%)	669 (68%)	751 (69%)
Count of tumour disease sites	2 (2 - 3)	3 (2 - 4)	2 (1 - 3)	3 (2 - 4)
Count of metastatic sites	2 (1 - 3)	2 (1 - 3)	2 (1 - 3)	2 (1 - 3)
Any prior trastuzumab all settings	1,430 (49%)	88 (11%)	991 (100%)	351 (32%)
Any prior lapatinib all settings	22 (1%)	0 (0%)	0 (0%)	22 (2%)
Any prior anthracycline all settings	1,403 (48%)	314 (39%)	605 (61%)	484 (44%)
Any prior taxane all settings	1,535 (53%)	185 (23%)	987 (100%)	363 (33%)
Lactate Dehydrogenase (U/L)				
Median (IQR)	267 (191-404)	277 (196-408)	244 (187-381)	284 (193-423)
Missing	119 (4%)	53 (7%)	43 (4%)	23 (2%)
Comorbidity count	2 (1 - 5)	2 (1 - 4)	3 (1 - 5)	3 (1 - 5)

Data are median (IQR) or number of patients (%).

Of the 2,894 patients, 402 were Randomised to receive Pertuzumab + Trastuzumab + Docetaxel (HTP), 406 to Placebo + Trastuzumab + Docetaxel (HT), 496 to Lapatinib + Capecitabine (LAPCAP), 495 to T-DM1 (T-DM1), 367 to Placebo + T-DM1 (T-DM1), 365 to Trastuzumab + Docetaxel/Paclitaxel (HT) and 363 to Pertuzumab + T-DM1 (T-DM1+P) (Table 23). Of the 2,894 patients, 46 did not have available adverse event follow-up (Table 23).

Table 25 presents the distribution of PROs within the pooled cohort according to study (missing data <10%). In the pooled cohort, 1,535 patients experienced grade ≥ 3 adverse events. Median follow-up was 50 months [95% CI: 49–51] in CLEOPATRA, 47 months [45–48] in EMILIA, and 54 months [54–55] in MARIANNE.

Table 25: Summary of pre-treatment patient-reported outcome values by study

	Total No. 2,894	CLEOPATRA No. 808	EMILIA No. 991	MARIANNE No. 1,095
Physical Well-being				
Median (IQR)	23 (18 - 26)	22 (18 - 26)	23 (18 - 26)	24 (19 - 27)
Missing	170 (6%)	57 (7%)	57 (6%)	56 (5%)
Social/family Well-being				
Median (IQR)	23 (18 - 26)	22 (18 - 26)	23 (19 - 26)	23 (19 - 26)
Missing	176 (6%)	59 (7%)	61 (6%)	56 (5%)
Emotional Well-being				
Median (IQR)	17 (13 - 20)	16 (12 - 19)	17 (13 - 20)	17 (13 - 19)
Missing	194 (7%)	63 (8%)	63 (6%)	68 (6%)
Functional Well-being				
Median (IQR)	18 (13 - 22)	17 (12 - 21)	18 (14 - 22)	18 (13 - 22)
Missing	145 (5%)	11 (1%)	65 (7%)	69 (6%)
Breast Cancer Subscale				
Median (IQR)	24 (20 - 28)	25 (21 - 29)	22 (18 - 25)	26 (21 - 30)
Missing	197 (7%)	69 (9%)	66 (7%)	62 (6%)
Trial Outcome Index score				
Median (IQR)	64 (54 - 73)	64 (54 - 74)	62 (52 - 71)	66 (56 - 76)
Missing	214 (7%)	69 (9%)	72 (7%)	73 (7%)
Total Fact B score				
Median (IQR)	102 (87 - 115)	101 (84 - 114)	100 (86 - 113)	104 (90 - 117)
Missing	219 (8%)	71 (9%)	74 (7%)	74 (7%)
Physical Well-being - by 10 unit				
Median (IQR)	2.3 (1.8 - 2.6)	2.2 (1.8 - 2.6)	2.3 (1.8 - 2.6)	2.4 (1.9 - 2.7)
Missing	170 (5.9%)	57 (7.1%)	57 (5.8%)	56 (5.1%)
Social/Family Well-being - by 10 unit				
Median (IQR)	2.3 (1.8 - 2.6)	2.2 (1.8 - 2.6)	2.3 (1.9 - 2.6)	2.3 (1.9 - 2.6)
Missing	176 (6.1%)	59 (7.3%)	61 (6.2%)	56 (5.1%)
Emotional Well-being - by 10 unit				
Median (IQR)	1.7 (1.3 - 2.0)	1.6 (1.2 - 1.9)	1.7 (1.3 - 2.0)	1.7 (1.3 - 1.9)
Missing	194 (6.7%)	63 (7.8%)	63 (6.4%)	68 (6.2%)
Functional Well-being - by 10 unit				
Median (IQR)	1.8 (1.3 - 2.2)	1.7 (1.2 - 2.1)	1.8 (1.4 - 2.2)	1.8 (1.3 - 2.2)
Missing	145 (5.0%)	11 (1.4%)	65 (6.6%)	69 (6.3%)
Breast Cancer Subscale - by 10 unit				
Median (IQR)	2.4 (2.0 - 2.8)	2.5 (2.1 - 2.9)	2.2 (1.8 - 2.5)	2.6 (2.1 - 3.0)
Missing	197 (6.8%)	69 (8.5%)	66 (6.7%)	62 (5.7%)
Trial Outcome Index score - by 10 unit				
Median (IQR)	6.4 (5.4 - 7.3)	6.4 (5.4 - 7.4)	6.2 (5.2 - 7.1)	6.6 (5.6 - 7.6)
Missing	214 (7.4%)	69 (8.5%)	72 (7.3%)	73 (6.7%)
Total Fact B score - by 10 unit				
Median (IQR)	10 (9 - 12)	10 (8 - 11)	10 (9 - 11)	10 (9 - 12)
Missing	219 (8%)	71 (9%)	74 (7%)	74 (7%)
Physical Well-being - by 10 unit				
[0.0,1.8]	574 (20%)	186 (23%)	204 (21%)	184 (17%)
[1.8,2.6]	1,584 (55%)	433 (54%)	559 (56%)	592 (54%)
[2.6,2.8]	566 (20%)	132 (16%)	171 (17%)	263 (24%)
Missing	170 (6%)	57 (7%)	57 (6%)	56 (5%)

Data are median (IQR) or number of patients (%).

Prognostic associations of PROs with survival outcomes

In the pooled cohort, the association between pre-treatment PROs and survival outcomes was best described by a linear association. The univariable and adjusted analysis identified significant associations for patient-reported physical well-being, trial outcome index score, total FACT-B score, functional well-being, and breast cancer subscale score with OS (Table 26).

Table 26: Univariable and adjusted association between patient-reported outcomes and overall survival

	Univariable					Adjusted#			
Patient-Reported Outcomes	**N**	**HR***	**95% CI**	**P-value**	***c***	**N**	**HR***	**95% CI**	**P-value**
Physical Well-being	2724	0.60	0.54-0.65	<0.001	0.60	2486	0.72	0.65-0.80	<0.001
Trial Outcome Index score	2680	0.83	0.80-0.86	<0.001	0.59	2449	0.87	0.84-0.91	<0.001
Total Fact-B score	2675	0.90	0.87-0.92	<0.001	0.57	2445	0.92	0.90-0.95	<0.001
Functional Well-being	2749	0.75	0.69-0.82	<0.001	0.56	2503	0.83	0.76-0.92	<0.001
Breast Cancer Subscale	2697	0.76	0.70-0.83	<0.001	0.55	2465	0.77	0.70-0.85	<0.001
Emotional Well-being	2700	0.86	0.78-0.96	0.008	0.52	2465	0.91	0.81-1.03	0.130
Social/Family Well-being	2718	0.96	0.88-1.05	0.383	0.51	2481	0.93	0.84-1.02	0.136

#Adjustment Variables: Sex, Age, Asian Race, ECOG PS, BMI, ER Status, PR Status, Time since initial diagnosis, Presence of visceral disease at baseline, Count of tumour disease sites, Any prior trastuzumab/anthracycline/taxane all settings, Lactate dehydrogenase at baseline and Comorbidity count
CI= Confidence interval, HR= Hazard Ratio, N = Number of subjects, *c* = concordance statistic, ECOG PS= Eastern cooperative oncology group performance status, BMI = Body Mass Index, and ER/PR = Estrogen Receptor/Progesterone Receptor
*HR based on 10-unit increase

Similar findings were also seen with PFS (Table 27).

Table 27: Univariable and adjusted association between patient-reported outcomes and progression-free survival

	Univariable					Adjusted#			
Patient-Reported Outcomes	**N**	**HR***	**95% CI**	**P-value**	***c***	**N**	**HR***	**95% CI**	**P-value**
Physical Well-being	2724	0.76	0.70-0.82	<0.001	0.55	2486	0.86	0.79-0.95	0.002
Trial Outcome Index score	2680	0.90	0.87-0.93	<0.001	0.55	2449	0.93	0.90-0.97	<0.001
Total Fact-B score	2675	0.93	0.91-0.95	<0.001	0.54	2445	0.95	0.93-0.98	<0.001
Functional Well-being	2749	0.84	0.78-0.90	<0.001	0.53	2503	0.90	0.83-0.98	0.011
Breast Cancer Subscale	2697	0.85	0.79-0.91	<0.001	0.53	2465	0.86	0.79-0.93	<0.001
Emotional Well-being	2700	0.88	0.81-0.97	0.010	0.52	2465	0.90	0.82-1.00	0.042
Social/Family Well-being	2718	0.93	0.86-1.01	0.075	0.51	2481	0.93	0.85-1.01	0.073

#Adjustment Variables: Sex, Age, Asian Race, ECOG PS, BMI, ER Status, PR Status, Time since initial diagnosis, Presence of visceral disease at baseline, Count of tumour disease sites, Any prior trastuzumab/anthracycline/taxane all settings, Lactate dehydrogenase at baseline and Comorbidity count
CI= Confidence interval, HR= Hazard Ratio, N = Number of subjects, *c* = concordance statistic, ECOG PS= Eastern cooperative oncology group performance status, BMI = Body Mass Index, and ER/PR = Estrogen Receptor/Progesterone Receptor
*HR based on 10-unit increase

Of the identified significant PROs, patient-reported physical well-being was the most prognostic PRO for OS (c=0.60) and PFS (c=0.55) (Table 26 & Table 27). Figure 9 presents the Kaplan-Meier estimates of survival outcomes by patient-reported physical well-being stratified by line of therapy.

Figure 9: Kaplan-Meier estimates of survival outcomes by patient-reported physical well-being and ECOG PS for participants initiating first and later line therapies

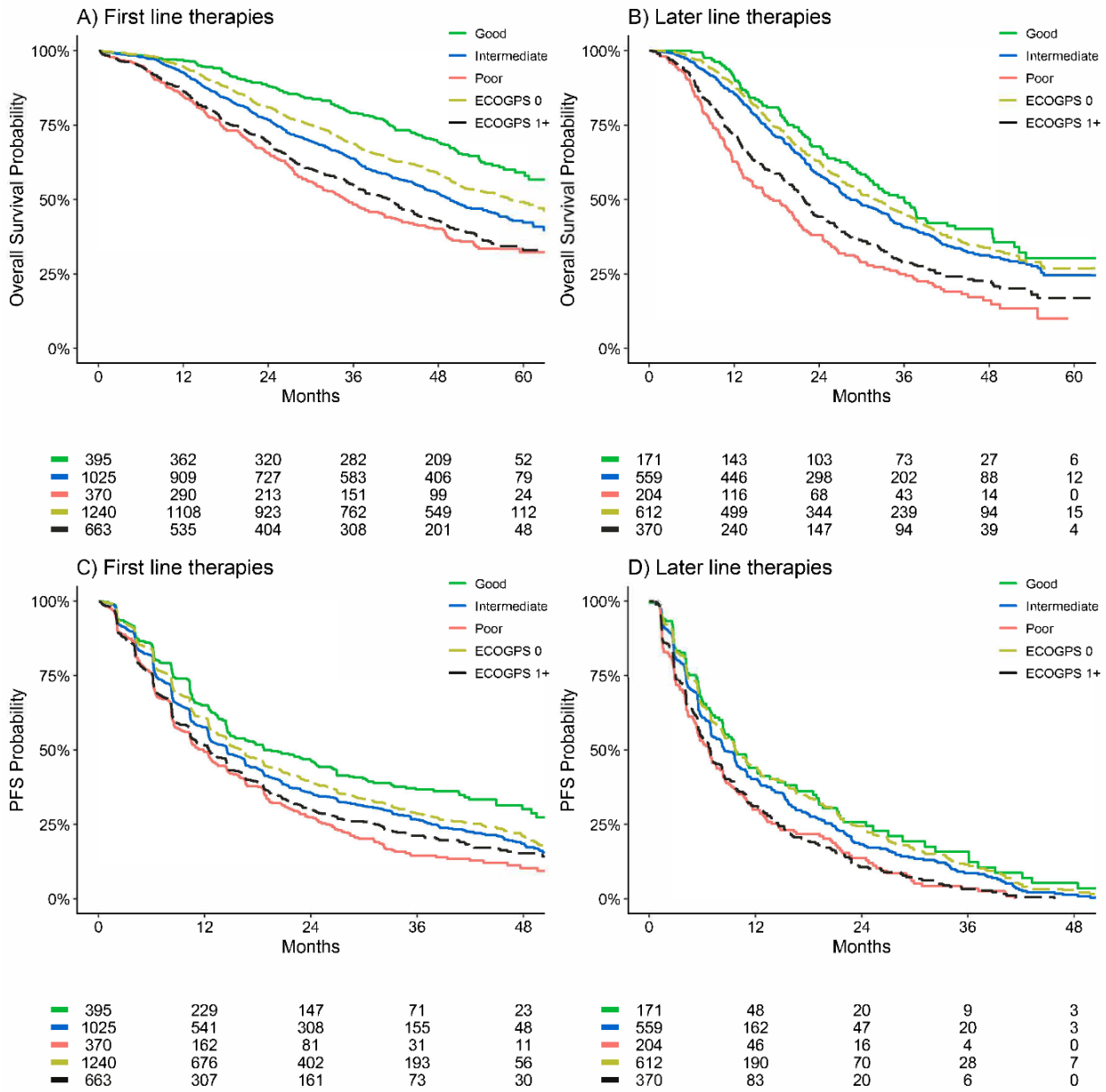


Figure 10 and Figure 11 presents the forest plots of the association between physical well-being and survival outcomes by clinical trial arms. Figure 9, Figure 10, and Figure 11 demonstrate that patients self-reporting “good” physical well-being had consistently improved survival outcomes compared to their counterparts who reported “poor” physical well-being, irrespective of treatment or line of therapy.

Figure 10: Forest plot presenting the association between physical well-being and overall survival by treatment arm

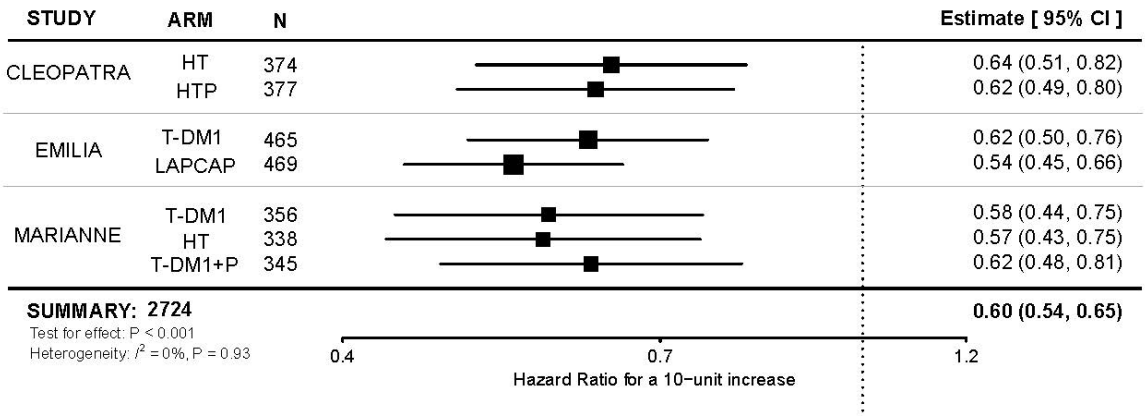
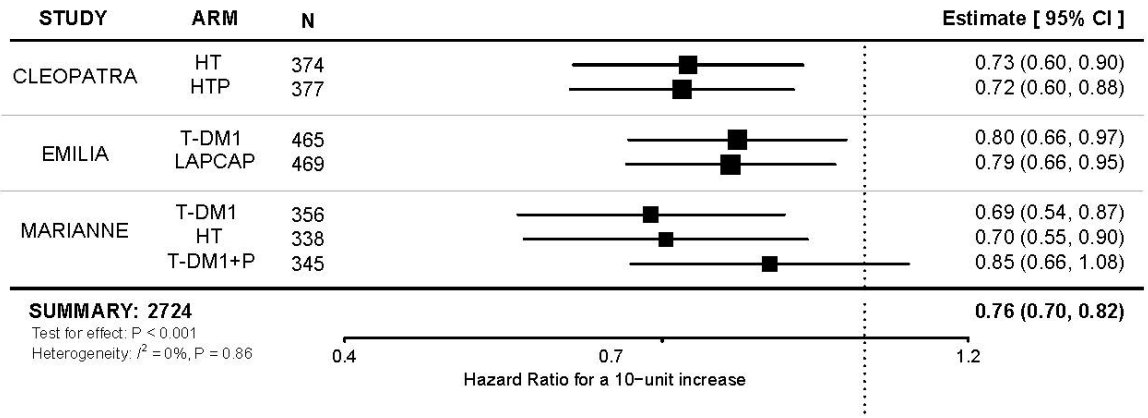


Figure 11: Forest plot presenting the association between physical well-being and progression free survival by treatment arm



Prognostic associations of PROs with grade ≥ 3 adverse events

Univariable and adjusted analysis identified significant associations for patient-reported physical well-being, trial outcome index score, total FACT-B score, functional well-being, and breast cancer subscale score with grade ≥ 3 adverse events (Table 28). Patient-reported physical well-being (*c* = 0.54) was the most prognostic.

Table 28: Univariable and adjusted associations between patient-reported outcomes and grade ≥ 3 adverse events

	Univariable					Adjusted[#]			
Patient-Reported Outcomes	**N**	**HR***	**95% CI**	**P-value**	***c***	**N**	**HR***	**95% CI**	**P-value**
Physical Well-being	2693	0.79	0.72-0.87	<0.001	0.54	2666	0.82	0.75-0.91	<0.001
Trial Outcome Index score	2650	0.92	0.89-0.96	<0.001	0.54	2624	0.94	0.90-0.97	<0.001
Total Fact-B score	2645	0.95	0.93-0.98	<0.001	0.54	2619	0.96	0.94-0.99	0.009
Functional Well-being	2718	0.88	0.81-0.96	0.003	0.53	2640	0.91	0.84-1.00	0.048
Breast Cancer Subscale	2666	0.88	0.81-0.96	0.004	0.52	2691	0.90	0.83-0.98	0.020
Emotional Well-being	2687	0.93	0.85-1.02	0.126	0.51	2660	0.98	0.89-1.07	0.622
Social/Family Well-being	2670	0.98	0.88-1.09	0.696	0.51	2643	0.96	0.86-1.07	0.465

#Adjustment Variables: Sex, Age, Asian Race, ECOG PS, BMI, ER Status, PR Status, Time since initial diagnosis, Presence of visceral disease at baseline, Count of tumour disease sites, Any prior trastuzumab/anthracycline/taxane all settings, Lactate dehydrogenase at baseline and Comorbidity count
CI= Confidence interval, HR= Hazard Ratio, N = Number of subjects, *c* = concordance statistic, ECOG PS= Eastern cooperative oncology group performance status, BMI = Body Mass Index, and ER/PR = Estrogen Receptor/Progesterone Receptor
*HR based on 10-unit increase

Figure 12 presents a Kaplan-Meier plot for the probability of developing grade ≥ 3 adverse events according to patient-reported physical well-being. presents a forest plot of the association between physical well-being and grade ≥ 3 adverse events by clinical trial arms. Figure 12 and Figure 13 demonstrate patients self-reporting “good” physical well-being consistently had less grade ≥ 3 adverse events compared to their counterparts who reported “poor” physical well-being, irrespective of treatment or line of therapy.

Figure 12: Kaplan-Meier estimates of grade 3 and above adverse events by patient-reported physical well-being and ECOG PS for participants initiating first and later line therapies

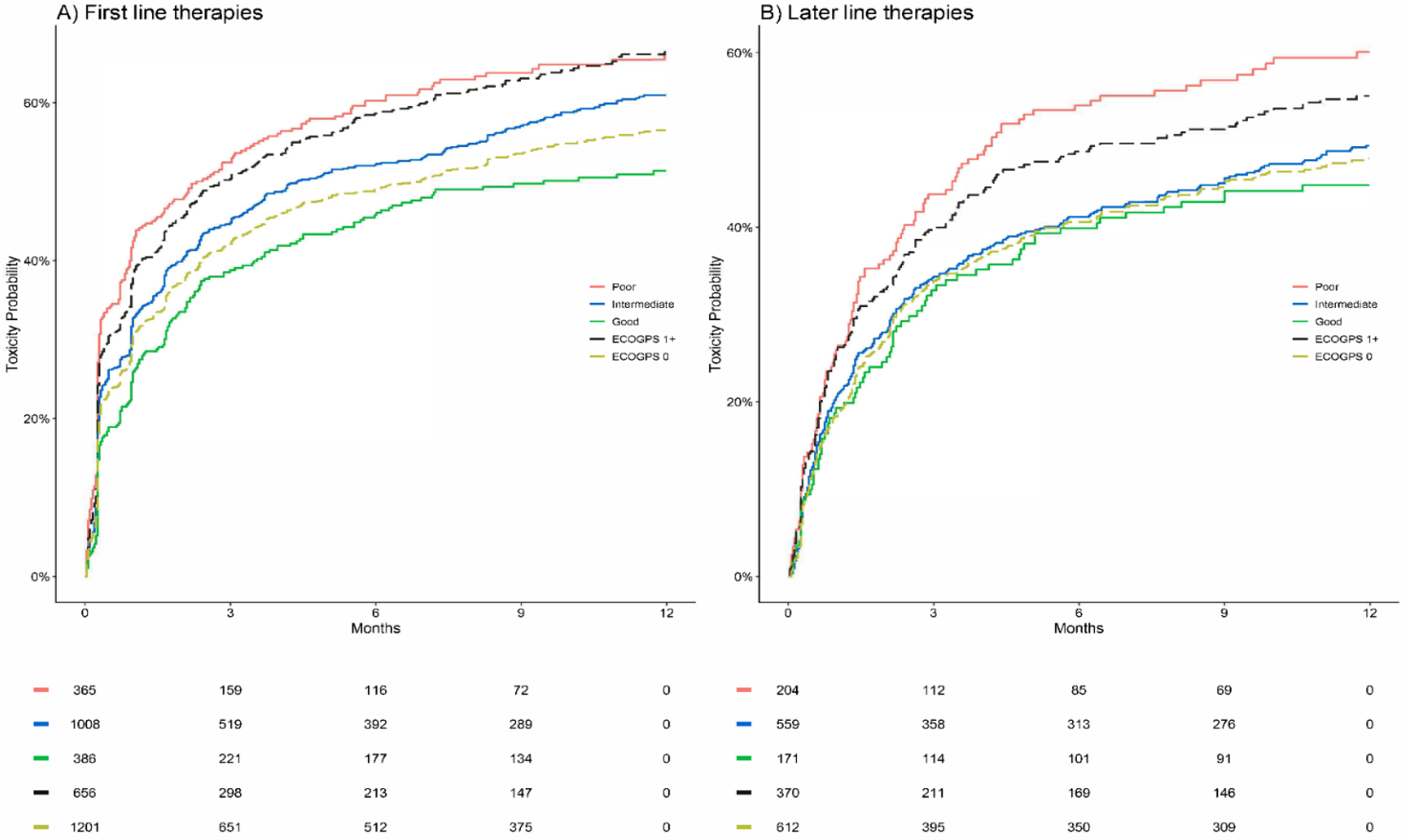
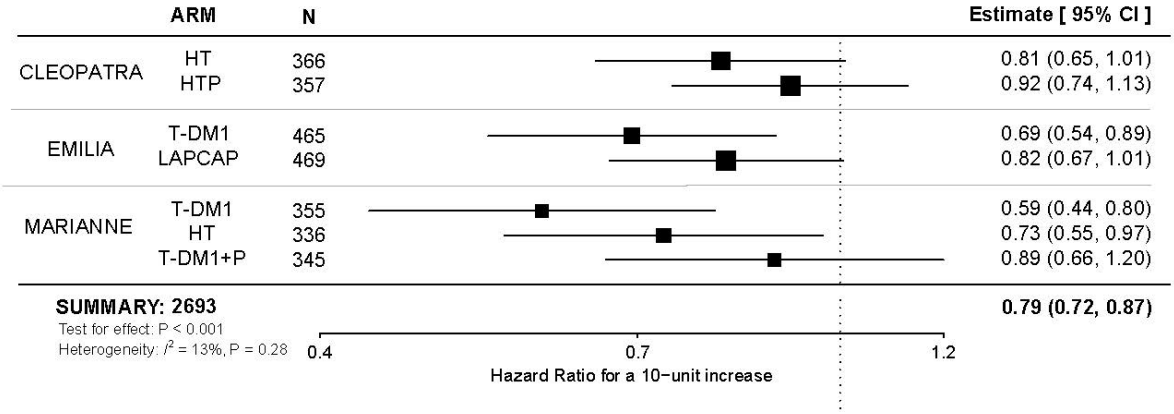


Figure 13: Forest plot presenting the association between physical well-being and grade ≥ 3 adverse events by treatment arm



Comparison of patient-reported physical well-being against ECOG PS

Of the 1,852 patients who had an ECOG PS score of 0 ('fully active, able to carry on all pre-disease performance without restriction'), 215 (12%) and 1,071 (58%) patients reported their physical well-being as "poor" or "intermediate", respectively (Table 29). Further, of the 1,852 patients with an ECOG PS 0, more than 25% specifically reported that they lacked energy and were in pain.

Table 29: Summary of physical well-being by ECOG PS score

	Total No. 2,885	0 No. 1,852	1+ No. 1,033
Physical Well-being Groups			
Poor	572 (20%)	215 (12%)	357 (35%)
Intermediate	1,583 (55%)	1,071 (58%)	512 (50%)
Good	563 (20%)	454 (25%)	109 (11%)
Missing	167 (6%)	112 (6%)	55 (5%)
I have a lack of energy			
Not at all	803 (28%)	638 (34%)	165 (16%)
A little bit	884 (31%)	566 (31%)	318 (31%)
Some-what	616 (21%)	362 (20%)	254 (25%)
Quite a bit	303 (11%)	134 (7%)	169 (16%)
Very much	107 (4%)	39 (2%)	68 (7%)
Missing	172 (6%)	113 (6%)	59 (6%)
I have nausea			
Not at all	1,990 (69%)	1,354 (73%)	636 (62%)
A little bit	449 (16%)	260 (14%)	189 (18%)
Some-what	172 (6%)	82 (4%)	90 (9%)
Quite a bit	72 (2%)	26 (1%)	46 (4%)
Very much	19 (1%)	8 (<1%)	11 (1%)
Missing	183 (6%)	122 (7%)	61 (6%)
Because of my physical condition, I have trouble meeting the needs of my family			
Not at all	1,275 (44%)	958 (52%)	317 (31%)
A little bit	663 (23%)	422 (23%)	241 (23%)
Some-what	407 (14%)	229 (12%)	178 (17%)
Quite a bit	225 (8%)	90 (5%)	135 (13%)
Very much	129 (4%)	34 (2%)	95 (9%)
Missing	186 (6%)	119 (6%)	67 (6%)
I have pain			
Not at all	859 (30%)	654 (35%)	205 (20%)
A little bit	826 (29%)	558 (30%)	268 (26%)
Some-what	492 (17%)	276 (15%)	216 (21%)
Quite a bit	358 (12%)	175 (9%)	183 (18%)
Very much	154 (5%)	59 (3%)	95 (9%)
Missing	196 (7%)	130 (7%)	66 (6%)
I am bothered by side effects of treatment			
Not at all	1,483 (51%)	1,009 (54%)	474 (46%)
A little bit	480 (17%)	314 (17%)	166 (16%)
Some-what	328 (11%)	179 (10%)	149 (14%)
Quite a bit	154 (5%)	79 (4%)	75 (7%)
Very much	68 (2%)	38 (2%)	30 (3%)
Missing	372 (13%)	233 (13%)	139 (13%)
I feel ill			
Not at all	1,291 (45%)	959 (52%)	332 (32%)
A little bit	699 (24%)	433 (23%)	266 (26%)
Some-what	405 (14%)	221 (12%)	184 (18%)
Quite a bit	211 (7%)	77 (4%)	134 (13%)
Very much	81 (3%)	30 (2%)	51 (5%)
Missing	198 (7%)	132 (7%)	66 (6%)
I am forced to spend time in bed			
Not at all	1,699 (59%)	1,244 (67%)	455 (44%)
A little bit	502 (17%)	286 (15%)	216 (21%)
Some-what	282 (10%)	129 (7%)	153 (15%)
Quite a bit	148 (5%)	42 (2%)	106 (10%)
Very much	59 (2%)	22 (1%)	37 (4%)
Missing	195 (7%)	129 (7%)	66 (6%)

Data are median (IQR) or number of patients (%)

On exploratory analysis, the OS prognostic performance (*c*) of patient-reported physical well-being (low vs. intermediate vs. high) in the pooled cohort was 0.58. Comparably, the OS prognostic performance of clinician-interpreted ECOG PS was 0.56 – this was statistically poorer than the patient-reported physical well-being groups (P<0.05) (Table 30). Nonetheless, on multivariable analysis, both physical well-being and ECOG PS remained statistically significant, indicating that both provide independent prognostic information (Table 30). Similar findings were also observed for PFS and grade ≥ 3 adverse events (Table 30). Additionally demonstrating the higher discrimination performance of patient-reported physical well-being compared to ECOG PS, the OS probability at 36 months in the "good" versus "poor" physical well-being groups ranged from 79% to 48%. Opposingly, the OS probability at 36 months for ECOG PS of 0 versus and 1+, ranged from 69% to 55% (Figure 9).

Table 30: Associations between patient-reported physical well-being and ECOG PS with overall survival, progression-free survival, and grade 3 and above adverse events

		Univariable				Multivariable#			
Predictors	**N**	**HR***	**[95% CI]**	**P-value**	***c***	**N**	**HR***	**[95% CI]**	**P-value**
Overall survival									
Physical Well-being				<0.001	0.58				<0.001
Good [a]	566	1				563	1		
Intermediate [b]	1584	1.50	[1.30-1.74]			1583	1.45	[1.25-1.68]	
Poor [c]	574	2.40	[2.03-2.84]			572	2.10	[1.76-2.49]	
ECOG PS				<0.001	0.56				<0.001
0	1852	1				1740	1		
1+	1033	1.59	[1.44-1.76]			978	1.39	[1.24-1.55]	
Progression-free survival									
Physical Well-being				<0.001	0.55				<0.001
Good [a]	566	1				563	1		
Intermediate [b]	1584	1.30	[1.15-1.47]			1583	1.27	[1.12-1.43]	
Poor [c]	574	1.73	[1.50-2.00]			572	1.59	[1.37-1.84]	
ECOG PS				<0.001	0.54				<0.001
0	1852	1				1740	1		
1+	1033	1.34	[1.22-1.47]			978	1.23	[1.11-1.35]	
Grade 3 and above adverse events									
Physical Well-being				<0.001	0.53				0.004
Good [a]	557	1				554	1		
Intermediate [b]	1567	1.17	[1.02-1.35]			1566	1.15	[1.00-1.32]	
Poor [c]	569	1.42	[1.21-1.67]			567	1.33	[1.12-1.57]	
ECOG PS				<0.001	0.53				<0.001
0	1813	1				1713	1		
1+	1026	1.30	[1.17-1.44]			974	1.22	[1.09-1.36]	

CI = confidence interval, ECOG PS = Eastern Cooperative Oncology Group Performance Status, HR = hazard ratio, *c* = concordance statistic.
*HR based on 10-unit increase
Model includes both pre-treatment physical well-being groups and ECOG PS
[a] Good physical well-being ≥ 2.6; [b] Intermediate physical well-being 1.8-2.59; [c] Poor physical well-being <1.8

Discussion

This study demonstrates for the first time, that pre-treatment PROs are significantly associated with OS, PFS, and grade ≥ 3 adverse events in patients with HER2-positive ABC treated with contemporary therapy. Additionally, this study found both patient-reported physical well-being and clinician-interpreted ECOG PS provide independent prognostic information.

Patient-reported physical well-being, trial outcome index score, total FACT-B score, functional well-being, and breast cancer subscale score were identified as significantly and independently associated with OS. Patient-reported physical well-being was the most prognostic PRO for primary and secondary outcomes. This is the first study to pool patients with HER2-positive ABC from three different trials that have been treated with contemporary therapies, and the results are consistent with recent findings in other advanced cancers [173-178].

The American Society of Clinical Oncology and the European Society for Medical Oncology (ESMO) has highlighted the identification of strategies that predict response and toxicity to anticancer therapies as key research priorities [188, 189]. In this study, we utilised PROs to identify patients with HER2-positive ABC who are more likely to achieve better survival outcomes and patients who are more likely to experience grade ≥ 3 adverse events. Routine and longitudinal collection of PROs in patients with advanced solid tumours treated with chemotherapy has been shown to improve quality of life, satisfaction, and survival outcomes [190-192]. PreCycle (NCT03220178) a multicentre, Randomised phase IV trial assessing the impact of longitudinally collected electronic PROs is showing positive preliminary results in patients with HR-positive/HER2-negative ABC [193]. Our study shows the potential value of PRO tools for facilitating shared decision-making and prognostic analysis in patients with HER2-positive ABC treated with a diverse range of anti-HER2 therapies. Therefore, we implore that the findings from this study are used to design strategies that bridge the gap between trials and routine clinical trials - as PROs are quite clearly prognostic of survival and toxicity for all the major contemporary treatments options in this ABC subtype.

At present PROs are primarily used in the oncology setting - as secondary outcomes of clinical trials to strengthen the interpretation of the primary outcomes (efficacy, safety, etc.) [170-172]. ESMO advocates the use of PROs as a co-primary endpoint in oncology trials, while the FDA is additionally advocating for their routine and standardised use as trial outcomes [194, 195]. Outside this, PROs are not used in oncology trial stratification and their clinical utility is only now emerging. Opposingly, clinician interpreted ECOG PS is routinely used to assess the eligibility of patients for clinical trials, as a prognostic factor for survival and toxicity outcomes in advanced cancers, and as an outcome measure [196, 197]. The present study demonstrates that patient-reported physical well-being has

independent, and potentially superior, prognostic performance to the clinician-interpreted ECOG PS. It is, therefore, essential that clinical practice transforms to place a greater emphasis on the patient's perspective and voice.

The findings of the present study are consistent with prior findings of patient-reported physical function/physical well-being and ECOG PS providing independent prognostic information [174, 177, 178]. It was interesting to note that 12% and 58% of patients classified as ECOG PS 0 reported poor and intermediate physical well-being, respectively. This indicates that 70% of the patients who were defined by their clinicians as 'fully active, and able to carry on all pre-disease performance without restrictions' reported limitations in their physical well-being status. The discordance between clinician-interpreted ECOG PS and patient-reported physical well-being suggests that appreciation of both parameters could allow for a more comprehensive prognostication of likely outcomes. Furthermore, it could be considered whether pre-treatment patient-reported physical well-being can be used as a stratification factor in clinical trials to optimise standardization between treatment arms.

RCTs are the backbone of evidence-based medicine, however, strict inclusion criteria within RCTs can limit the generalizability of results (for example, the study cohort was almost entirely restricted to participants with an ECOG PS of 1 or less) [155]. It is also acknowledged that some PROs data were missing – as some patients may not have answered the FACT-B questionnaire at baseline. However, RCTs provide a rigorous, high-quality collection of patient-reported outcomes, survival outcomes, and adverse event data [156]. Additionally, this study pooled large (n=2,894) data from three trials (CLEOPATRA, EMILIA, and MARIANNE) to increase study power and generalizability. The availability of such detailed data was essential in accurately assessing the relationship between ECOG PS and patient-reported physical well-being in HER2-positive ABC patients, illustrating how IPD sharing enhances the applicability of trial findings to broader patient populations. Effective communication is a core component of shared decision-making [198] and can be enhanced with the use of patient-reported questionnaires that incorporate health-related quality of life measures as well as clinically interpreted measures. Future research should examine the association between ECOG PS and PROs in early breast cancer, other breast cancer subtypes, and in real-world populations – which are more likely to have broader distributions of ECOG PS and PROs scores.

In conclusion, pre-treatment PROs had a significant relationship with both survival and toxicity outcomes in patients with HER2-positive ABC, initiating contemporary anticancer treatment. Additionally, patient-reported physical well-being and clinician-interpreted ECOG PS were found to provide independent prognostic information. The study highlights the potential of combining patient-reported questionnaires and clinically interpreted measures to enhance clinical trial design and provide clinical insights that facilitate shared decision-making in breast cancer.

CHAPTER 8: SIGNIFICANCE, FUTURE DIRECTIONS AND CONCLUSION

The significance of this thesis lies in its potential to profoundly impact both precision medicine and the broader landscape of data transparency. Over the last two decades, data transparency has evolved from a niche concern to a global priority. It is driven by the increasing need for openness in scientific inquiry and the recognition of patients' rights to access information from the clinical trials in which they participate [26]. Historically, regulatory frameworks focused primarily on the confidentiality of clinical trial data, but initiatives like the *AllTrials* campaign have highlighted the necessity for full disclosure of clinical trial results [199]. These shifts have sparked important discussions about the ethics, feasibility, and societal impact of transparent data sharing.

This thesis addresses both the challenges and the opportunities presented by these shifts in transparency policies. It explores the practical applications of transparent data sharing, particularly in clinical settings, where access to trial data can improve patient care and accelerate scientific progress. Ensuring that clinical trial data is accessible holds the potential to overcome research barriers, especially in low-resource settings where large-scale trials are often unfeasible. Shared data can bridge this gap. Researchers in such regions can use the shared data to build on existing datasets and contribute to advancements in medical research. In turn, this promotes a more equitable global healthcare system by widening research opportunities and reducing disparities in access to information.

The importance of this work can be understood through two major themes: enhancing data transparency and its practical applications in clinical settings.

Enhancing Data Transparency

Transparent sharing of clinical trial data can foster greater scientific integrity, enhance patient safety, and accelerate medical research. Although data sharing policies are designed to increase the accessibility of clinical trial data, challenges persist in translating these policies into effective and widespread practices. Despite notable commitments by the pharmaceutical industry, through PhRMA and EFPIA, to share anonymised IPD from clinical trials for approved medicines with qualified researchers, significant gaps remain [14, 43].

This thesis reveals that vast amounts of valuable data from pivotal clinical trials conducted in the prior decade remain inaccessible. These barriers are particularly evident in oncology trials, where public access to data could drive significant progress in understanding treatment efficacy and patient

outcomes. Despite growing recognition of the benefits of transparency, the lack of standardised global frameworks and inconsistencies in data sharing practices across the pharmaceutical industry continue to hinder efforts to create a unified approach. Moreover, technological and logistical challenges further complicate the effective implementation of data sharing policies. Variations in data formats, inadequate anonymisation techniques, and the absence of centralised data repositories make it difficult for researchers to aggregate and utilise data efficiently [14]. Addressing these challenges requires not only stronger policy enforcement but also the development of global standards that promote seamless integration of clinical trial data across platforms. These changes would enhance the ability of researchers to conduct large-scale meta-analyses and secondary analysis.

The audit of IPD availability for anticancer and high revenue-generating medicines represents a significant step forward in understanding IPD sharing practices for pivotal trials. IPD was accessible for 45% of oncology trials linked to FDA-approved drug labels from 2011 to 2021, whereas IPD was accessible for 64% of clinical trials supporting the FDA approval of the top 30 revenue-generating medicines in 2021. These figures indicate notable progress but also highlight the high risk of bias when relying on this limited IPD for systematic reviews and meta-analysis [53]. A notable case illustrating the risks associated with inadequate data sharing is the controversy surrounding the drug Tamiflu [200]. The lack of access to comprehensive trial data, including IPD, CSRs and detailed summaries, hindered independent assessment of the medicine efficacy and safety profile [201]. Transparent sharing of CSRs and IPD could have mitigated some of the issues by allowing for thorough, unbiased analyses. While other factors also played significant roles, this case highlighted how comprehensive data transparency – sharing of CSRs and IPD - is needed.

The examination of industry-wide trends in data transparency highlights the need for comprehensive and enforceable policies. The growing volume of data generated during clinical trials necessitates stronger regulations to ensure that this wealth of information is made accessible to the broader scientific community. In addition to promoting scientific progress, accessible data plays an important role in fulfilling the ethical obligations owed to clinical trial participants. These individuals often enrol in clinical trials with the understanding that their involvement, while potentially not personally beneficial, will contribute to advancing future patient care [26]. The ethical dimension of data transparency extends beyond access. It also concerns how informed consent is structured. Informed consent models must evolve to reflect the changing landscape of clinical research [202]. Rather than creating additional barriers, informed consent should be designed to support the goal of knowledge generation, empowering participants with clear information about how their data will contribute to broader scientific and healthcare improvements. This approach respects participants' contributions

while reinforcing the essential purpose of trials in advancing medical knowledge for future generations

The findings of this thesis show that the industry's early concerns regarding patient privacy and financial costs are no longer major barriers. Instead, the variability in data sharing policies and lack of clear regulatory guidelines are the main challenges [14, 43]. Through a detailed analysis of data sharing practices for clinical trials supporting regulatory approvals, this thesis provides actionable recommendations for improving data sharing rates and practices.

Firstly, the adoption of open-access IPD sharing models or the management of IPD sharing processes by independent parties is important. When data sharing is overseen by an independent entity rather than the sponsoring company or party with vested interests, the risk of bias is significantly reduced. Independent oversight minimises conflicts of interest and thus ensuring that the sharing and analysis of clinical trial data are driven solely by scientific and ethical considerations. This approach promotes greater confidence in the data, as researchers and the public can trust that decisions regarding data accessibility are objective. Secondly, it is essential to ensure that IPD is immediately eligible for sharing upon the medicine’s approval. Delays in making IPD available often results in missed opportunities for early reassessment of clinical trial outcomes, slowing the pace of scientific discovery and innovation. This issue is particularly pertinent when extended follow-up periods create unjustifiable barriers to data access, potentially withholding important safety and efficacy information from independent researchers. Therefore, standardised policies that stipulate the immediate availability of all IPD underlying the results presented in drug labels are necessary. Such policies would ensure that data supporting the approval of medicines are promptly available for independent scrutiny and scientific investigation, addressing the concern that extended follow-up periods often delay or deny access to trial data, undermining the purpose of data sharing policies.

Practical Applications in Clinical Settings

Enhanced data transparency, particularly through the sharing of IPD is important for advancing medical research. The development of clinical prediction models for adverse effects such as diarrhoea and neutropenia following abemaciclib treatment showcases how granular data - such as IPD - can inform tailored interventions. These prediction models enable clinicians to understand which patients are at higher risk for certain side effects, allowing for proactive risk management. This level of personalization aligns with the principles of precision medicine, optimising patient care by ensuring that treatments are both safe and effective, minimising unnecessary suffering, and improving patient outcomes. Furthermore, precision medicine's focus on data-driven decisions means that healthcare is moving toward an era where treatments are not just standardised across

populations but are tailored specifically to an individual's unique genetic, clinical, and lifestyle factors. Transparent sharing of IPD is important in achieving this vision since it provides the necessary data to refine treatment protocols and identify predictors of both positive and adverse treatment responses.

The assessment of the prognostic performance of pre-treatment PROs in predicting prognosis and toxicity for patients receiving contemporary treatments for HER2-positive advanced breast cancer further illustrates the practical applications of IPD transparency. Integrating PROs into clinical practice can enhance shared decision-making between clinicians and patients since PROs provide a comprehensive view of a patient's health status and likely treatment outcomes. This integration represents a significant advancement in precision medicine, where the patient's voice and experiences are integral to the treatment process, ultimately leading to patient-centred care. The practical applications of IPD sharing in clinical settings are vast and varied. Whether in the development of predictive models for treatment side effects or the integration of PROs to guide patient-centred care, transparent sharing of IPD plays an important role in advancing precision medicine. As data sharing initiatives become more widespread, and as clinicians and researchers continue to develop tools that make use of this rich data, the potential for more accurate, personalised, and compassionate care will only continue to grow. Therefore, chapters 6 and 7 complete the cycle outlined at the beginning of this thesis, transitioning from the initial audit of IPD availability and advocacy for enhanced transparency to showcasing the practical applications of data sharing.

Future Directions

Building on the findings of this thesis, several key initiatives should be prioritised to further advance the field of data transparency and its practical applications in healthcare:

1) Conducting surveys to gather and validate the perspectives of clinical trial participants on data sharing is essential. Incorporating participant voices in data sharing policies can help align these policies with participant expectations, emphasising transparency as a shared goal.
2) Regulatory bodies should work towards standardising data sharing policies across therapeutic areas and jurisdictions. Currently, the variability in data sharing requirements and regulations creates significant barriers for accessing and utilising clinical trial data. Harmonisation of these policies will reduce this variability and ensure uniform and equitable access to IPD worldwide.

3) Establishing centralised repositories that house IPD from clinical trials would be a major step forward in streamlining access to valuable data. Such repositories would provide a single, organised platform where researchers can easily access IPD for meta-analyses, secondary analyses, and validation studies. Ensuring that these repositories focus on secure and accessible storage - without unnecessary data redaction or over anonymization - will maximise their usefulness and allow researchers to identify patterns and trends that are often obscured in smaller, isolated datasets.
4) One of the key challenges in making data widely available is the need to protect patient privacy. As data sharing becomes more widespread, improving anonymisation techniques will be the key in addressing these privacy concerns. Development of more sophisticated methods of de-identifying patient data will allow us to balance the need for privacy with the benefits of open data access. Therefore, it is hoped that enhanced anonymisation will ensure that patient identities are protected while still allowing the data to be useful for research. This means that there will be a greater degree of openness without compromising ethical standards.
5) Promoting collaboration between pharmaceutical companies, academia, regulatory bodies, and patient advocacy groups will create a cohesive data sharing ecosystem. Each of these stakeholder groups bring a unique perspective and set of resources to the table. When they work together, they can drive the development of shared goals and frameworks that support data transparency.
6) Integrating real-world data from electronic health records, registries, and other sources with clinical trial data can enhance the generalizability of findings. This integration would also allow for more accurate predictions of treatment outcomes across a broader spectrum of patients, supporting the development of more effective and personalised treatments.
7) Conducting studies to assess the impact of data sharing initiatives on clinical outcomes will demonstrate their tangible benefits and thus encourage wider adoption of data sharing practices.

To advance the initiatives highlighted in this thesis, the initial step will involve conducting patient surveys to gather contemporary perspectives on data sharing. Engaging directly with clinical trial participants will allow us to gain valuable insights into their expectations, concerns, and willingness to share their data. This approach will refresh and expand upon previous work, ensuring that patient voices remain central in shaping data sharing policies. Following this, a large-scale audit of IPD availability across different therapeutic areas will map the current landscape. This audit will help in identifying both strengths and gaps in accessibility and transparency. Pinpointing gaps and inconsistencies will help reveal the barriers independent researchers face in accessing vital

information. For clinical trials deemed eligible for independent IPD requests, an evaluation of data accessibility, quality, and overall utility will assess whether shared data meets standards for secondary analysis. This evaluation will also consider whether anonymisation techniques effectively balance data usability and privacy. Finally, accessible data will be leveraged in large-scale meta-analyses aimed to uncover practice-changing insights.

These meta-analyses are expected to provide substantial evidence to advocate for regulatory mandates on IPD availability, emphasising the necessity for standardised and enforceable policies that ensure immediate eligibility of IPD for sharing upon a medicine's approval. Meta-analyses based on shared IPD will also offer a significant opportunity to address the limitations of aggregated data. It is expected that they will reveal patient-specific characteristics and treatment responses that are often masked in broader datasets. For example, while aggregated data may show the efficacy of a treatment in a general population, meta-analyses of IPD can identify whether specific subgroups of patients - such as those with particular genetic markers or comorbidities - respond differently. These individualised insights are essential for the advancement of precision medicine. They also help to reduce the risk of applying broad, less effective treatment strategies across diverse populations.

This drive for greater transparency must be coupled with a commitment to improving the ethical and operational frameworks that govern data sharing. Regulatory bodies must not only enforce policies but also foster an ecosystem where shared data translates into tangible clinical outcomes. This is where the future of research lies: in creating a collaborative environment where data is not siloed but used to fuel innovation, improve patient care, and advance the principles of precision medicine.

In conclusion, this thesis highlights its significance in advancing both precision medicine and data transparency. These are the two key areas poised to reshape the landscape of modern healthcare. The work presented in this thesis addresses pressing global challenges in data sharing, focusing on the ethical, operational, and regulatory hurdles that continue to impede full transparency. Through a critical examination of current practices, particularly in the context of clinical trials and their impact on patient outcomes, this thesis offers clear recommendations to enhance the accessibility and utility of clinical trial data. Advocating for standardised global frameworks, independent oversight of data sharing, and immediate access to IPD post-approval, my work has highlighted the vital role of transparent data in promoting scientific integrity and patient-centred care. Integration of such data into clinical practice, as shown through the development of prediction models for treatment side effects and the inclusion of PROs, demonstrates the practical applications of transparency in achieving personalised treatment plans and improving patient outcomes. Ongoing collaboration between pharmaceutical companies, academia, and patient advocacy groups will create a unified and a sustainable approach. Ultimately, this thesis has endeavoured to restore the core principles

of scientific inquiry: openness, collaboration, and patient-centredness. This pursuit of greater transparency honours the contributions of clinical trial participants and paves the way for transformative advancements in precision medicine and patient care. With each discovery, we hope to gain a more profound understanding of our origins in medical research.

CHAPTER 9: BIBLIOGRAPHY

1. Umscheid, C.A., D.J. Margolis, and C.E. Grossman, *Key concepts of clinical trials: a narrative review.* Postgraduate medicine, 2011. **123**(5): p. 194-204.
2. Institute of Medicine, *Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risk*. 2015, Washington, DC: The National Academies Press. 304.
3. Rydzewska, L.H.M., L.A. Stewart, and J.F. Tierney, *Sharing individual participant data: through a systematic reviewer lens.* Trials, 2022. **23**(1): p. 167.
4. Hopkins, A.M., et al., *Development and Validation of a Prognostic Model for Patients with Advanced Lung Cancer Treated with the Immune Checkpoint Inhibitor Atezolizumab.* Clinical Cancer Research, 2020. **26**(13): p. 3280-3286.
5. Hopkins, A.M., et al., *Concomitant Antibiotic Use and Survival in Urothelial Carcinoma Treated with Atezolizumab.* European Urology, 2020. **78**(4): p. 540-543.
6. Thibault, R.T., et al., *Open Science 2.0: Towards a truly collaborative research ecosystem.* PLoS Biol, 2023. **21**(10): p. e3002362.
7. Shahin, M.H., et al., *Open Data Revolution in Clinical Research: Opportunities and Challenges.* Clinical and Translational Science, 2020. **13**(4): p. 665-674.
8. Zarin, D.A. and T. Tse, *Medicine. Moving toward transparency of clinical trials.* Science, 2008. **319**(5868): p. 1340-2.
9. DeVito, N.J., et al., *Availability of results of clinical trials registered on EU Clinical Trials Register: cross sectional audit study.* BMJ Med, 2024. **3**(1): p. e000738.
10. Bradley, S.H., et al., *Reducing bias and improving transparency in medical research: a critical overview of the problems, progress and suggested next steps.* J R Soc Med, 2020. **113**(11): p. 433-443.
11. Hopkins, A.M., A. Rowland, and M.J. Sorich, *Data sharing from pharmaceutical industry sponsored clinical studies: audit of data availability.* BMC Medicine, 2018. **16**(1): p. 165.
12. Miller, J., et al., *Sharing of clinical trial data and results reporting practices among large pharmaceutical companies: cross sectional descriptive study and pilot of a tool to improve company practices.* BMJ, 2019. **366**: p. l4217.
13. PhRMA and EFPIA. *Principles for responsible clinical trial data sharing*. 2013 2021]; Available from: https://www.efpia.eu/media/25666/principles-for-responsible-clinical-trial-data-sharing.pdf.
14. Modi, N.D., et al., *A 10-year update to the principles for clinical trial data sharing by pharmaceutical companies: perspectives based on a decade of literature and policies.* BMC Medicine, 2023. **21**(1): p. 400.
15. Vivli. *Center for Global Clinical Research Data*. 2021 [cited 2021 November 5]; Available from: https://vivli.org/.
16. ClinicalStudyDataRequest. 2021 [cited 2021 October 24]; Available from: https://www.clinicalstudydatarequest.com/.
17. Rockhold, F., P. Nisen, and A. Freeman, *Data Sharing at a Crossroads.* New England Journal of Medicine, 2016. **375**(12): p. 1115-1117.
18. Hopkins, A.M., et al., *Heterogeneity and Utility of Pharmaceutical Company Sharing of Individual-Participant Data Packages.* JAMA Oncology, 2023. **9**(12): p. 1621-1626.
19. Tucker, K., et al., *Protecting patient privacy when sharing patient-level data from clinical trials.* BMC Medical Research Methodology, 2016. **16**(1): p. 77.
20. Hut-Mossel, L., et al., *Understanding how and why audits work in improving the quality of hospital care: A systematic realist review.* PLOS ONE, 2021. **16**(3): p. e0248677.
21. Bush, J., *Why Auditing Is Important*, in *Practical Database Auditing for Microsoft SQL Server and Azure SQL: Troubleshooting, Regulatory Compliance, and Governance*, J. Bush, Editor. 2022, Apress: Berkeley, CA. p. 3-7.
22. Bonini, S., et al., *Transparency and the European Medicines Agency — Sharing of Clinical Trial Data.* New England Journal of Medicine, 2014. **371**(26): p. 2452-2455.

23. Fogel, D.B., *Factors associated with clinical trials that fail and opportunities for improving the likelihood of success: A review.* Contemporary clinical trials communications, 2018. **11**: p. 156-164.
24. Doshi, P., F. Godlee, and K. Abbasi, *Covid-19 vaccines and treatments: we must have raw data, now.* BMJ : British Medical Journal, 2022. **376**: p. o102.
25. El Emam, K., S. Rodgers, and B. Malin, *Anonymising and sharing individual patient data.* BMJ : British Medical Journal, 2015. **350**: p. h1139.
26. Mello, M.M., V. Lieou, and S.N. Goodman, *Clinical Trial Participants' Views of the Risks and Benefits of Data Sharing.* New England Journal of Medicine, 2018. **378**(23): p. 2202-2211.
27. Geifman, N., et al., *Opening clinical trial data: are the voluntary data-sharing portals enough?* BMC Medicine, 2015. **13**(1): p. 280.
28. Lo, B., *Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risk.* JAMA, 2015. **313**(8): p. 793-794.
29. Loder, E., *Data sharing: making good on promises.* BMJ : British Medical Journal, 2018. **360**: p. k710.
30. EFPIA. *Pharmaceutical companies continue to drive transparency and underline industry investment in Europe's Healthcare*. 2017 [cited 2021 November 5]; Available from: https://www.efpia.eu/news-events/the-efpia-view/statements-press-releases/20062018-pharmaceutical-companies-continue-to-drive-transparency-and-underline-industry-investment-in-europe-s-healthcare/.
31. Yale School of Medicine. *YODA*. 2021 [cited 2021 November 5]; Available from: https://yoda.yale.edu/.
32. Ohmann, C., et al., *Sharing and reuse of individual participant data from clinical trials: principles and recommendations.* BMJ Open, 2017. **7**(12): p. e018647.
33. The International Consortium of Investigators for Fairness in Trial Data Sharing, *Toward Fairness in Data Sharing.* New England Journal of Medicine, 2016. **375**(5): p. 405-407.
34. Strom, B.L., et al., *Data Sharing — Is the Juice Worth the Squeeze?* New England Journal of Medicine, 2016. **375**(17): p. 1608-1609.
35. National Cancer Institute. *Cancer.Gov - A to Z List of Cancer Drugs*. 2021 [cited 2021 October 24]; Available from: https://www.cancer.gov/about-cancer/treatment/drugs.
36. Mullard, A. *2020 FDA Drug Approvals*. FDA Drug Approvals 2021 [cited 2021 October 20]; Available from: https://www.nature.com/articles/d41573-021-00002-0.
37. National Library of Medicine. *DailyMed*. 2021 [cited 2021 October 28]; Available from: https://dailymed.nlm.nih.gov/dailymed/.
38. Urquhart, L. *Top companies and drugs by sales in 2020*. Top companies and drugs by sales 2021 [cited 2021 January 7]; Available from: https://www.nature.com/articles/d41573-021-00050-6.
39. Fierce Pharma. *The top 20 pharma companies by 2020 revenue*. 2021 [cited 2021 October 24]; Available from: https://www.fiercepharma.com/special-report/top-20-pharma-companies-by-2020-revenue.
40. S&P Global Market Intelligence. *Cancer drugs show big sales rise in 2020*. 2021 [cited 2021 November 5]; Available from: https://www.spglobal.com/marketintelligence/en/news-insights/latest-news-headlines/cancer-drugs-show-big-sales-rise-in-2020-abbvie-s-humira-again-tops-list-63249898.
41. Buntz, B. *50 of 2020's best-selling pharmaceuticals*. 2021 [cited 2021 October 24]; Available from: https://www.drugdiscoverytrends.com/50-of-2020s-best-selling-pharmaceuticals/.
42. R Core Team. *R: A language and environment for statistical computing. R Foundation for Statistical Computing*. 2023 [cited 2023 January 20]; Available from: https://www.R-project.org/.
43. Modi, N.D., et al., *Audit of Data Sharing by Pharmaceutical Companies for Anticancer Medicines Approved by the US Food and Drug Administration.* JAMA Oncology, 2022. **8**(9): p. 1310-1316.

44. Murugiah, K., et al., *Availability of Clinical Trial Data From Industry-Sponsored Cardiovascular Trials.* Journal of the American Heart Association, 2016. **5**(4): p. e003307.
45. Goldacre, B., et al., *Pharmaceutical companies' policies on access to trial data, results, and methods: audit study.* BMJ : British Medical Journal, 2017. **358**: p. j3334.
46. Axson, S.A., et al., *Clinical trial transparency and data sharing among biopharmaceutical companies and the role of company size, location and product type: a cross-sectional descriptive analysis.* BMJ Open, 2021. **11**(7): p. e053248.
47. Cochrane Library. *A statement in support of EMA's clinical study report transparency policy*. 2020 [cited 2023 December 23]; Available from: https://www.cochrane.org/news/statement-support-emas-clinical-study-report-transparency-policy.
48. International Committee of Medical Journal Editors. *Recommendations*. 2024 [cited 2024 January 2]; Available from: https://www.icmje.org/recommendations/.
49. World Health Organization. *Joint statement on public disclosure of results from clinical trials*. 2017 [cited 2023 December 23]; Available from: https://www.who.int/news/item/18-05-2017-joint-statement-on-registration.
50. Fallah, J., et al., *Survival outcomes in older men with non-metastatic castration-resistant prostate cancer treated with androgen receptor inhibitors: a US Food and Drug Administration pooled analysis of patient-level data from three randomised trials.* Lancet Oncol, 2021. **22**(9): p. 1230-1239.
51. Taylor, C., et al., *Radiotherapy to regional nodes in early breast cancer: an individual patient data meta-analysis of 14,324 women in 16 trials.* The Lancet, 2023. **402**(10416): p. 1991-2003.
52. Bradley, R., et al., *Aromatase inhibitors versus tamoxifen in premenopausal women with oestrogen receptor-positive early-stage breast cancer treated with ovarian suppression: a patient-level meta-analysis of 7030 women from four randomised trials.* The Lancet Oncology, 2022. **23**(3): p. 382-392.
53. Page, M.J., J.P.T. Higgins, and J.A.C. Sterne, *Assessing risk of bias due to missing results in a synthesis*, in *Cochrane Handbook for Systematic Reviews of Interventions*. 2019. p. 349-374.
54. Stewart, L.A. and M.J. Clarke, *Practical methodology of meta-analyses (overviews) using updated individual patient data. Cochrane Working Group.* Stat Med, 1995. **14**(19): p. 2057-79.
55. European Medicines Agency. *Clinical data publication*. 2024 [cited 2024 January 23]; Available from: https://www.ema.europa.eu/en/human-regulatory/marketing-authorisation/clinical-data-publication.
56. Food and Drug Administration. *FDA Continues to Support Transparency and Collaboration in Drug Approval Process as the Clinical Data Summary Pilot Concludes*. 2020 [cited 2023 December 23]; Available from: https://www.fda.gov/news-events/press-announcements/fda-continues-support-transparency-and-collaboration-drug-approval-process-clinical-data-summary.
57. Health Canada. *Public Release of Clinical Information: consultation*. 2019 [cited 2023 November 19]; Available from: https://www.canada.ca/en/health-canada/programs/consultation-public-release-clinical-information-drug-submissions-medical-device-applications.html.
58. National Institute of Health. *Which Policies apply to my Research?* 2024 [cited 2024 January 2]; Available from: https://sharing.nih.gov/other-sharing-policies/which-policies-apply-to-my-research.
59. Hopkins, A.M., et al., *Accessibility of clinical study reports supporting medicine approvals: a cross-sectional evaluation.* Journal of Clinical Epidemiology, 2024.
60. National Academies of Sciences and Engineering and Medicine, *The National Academies Collection: Reports funded by National Institutes of Health*, in *Reflections on Sharing Clinical Trial Data: Challenges and a Way Forward: Proceedings of a Workshop*, C. Shore, et al., Editors. 2020, National Academies Press (US): Washington (DC).

61. National Institute of Health. *Data Management and Sharing Policy*. 2024 [cited 2024 April 7]; Available from: https://sharing.nih.gov/data-management-and-sharing-policy.
62. Menikoff, J., J. Kaneshiro, and I. Pritchard, *The Common Rule, Updated.* New England Journal of Medicine, 2017. **376**(7): p. 613-615.
63. Bauchner, H., R.M. Golub, and P.B. Fontanarosa, *Data Sharing: An Ethical and Scientific Imperative.* JAMA, 2016. **315**(12): p. 1238-1240.
64. Taichman, D.B., et al., *Data Sharing Statements for Clinical Trials: A Requirement of the International Committee of Medical Journal Editors.* JAMA, 2017. **317**(24): p. 2491-2492.
65. Zemła-Pacud, Ż. and G. Lenarczyk, *Clinical Trial Data Transparency in the EU: Is the New Clinical Trials Regulation a Game-Changer?* IIC - International Review of Intellectual Property and Competition Law, 2023. **54**(5): p. 732-763.
66. Taichman, D.B., et al., *Sharing Clinical Trial Data — A Proposal from the International Committee of Medical Journal Editors.* New England Journal of Medicine, 2016. **374**(4): p. 384-386.
67. Warren, E., *Strengthening Research through Data Sharing.* New England Journal of Medicine, 2016. **375**(5): p. 401-403.
68. Bioethics International. *Good Pharma Scorecard*. 2023 [cited 2023 July 20]; Available from: https://bioethicsinternational.org/good-pharma-scorecard/.
69. Buntz, B. *50 of 2021's best-selling pharmaceuticals*. 2022 [cited 2022 September 20]; Available from: https://www.drugdiscoverytrends.com/50-of-2021s-best-selling-pharmaceuticals/.
70. Gibney Michael, J.H. *Pfizer's COVID-19 shot is 2021's best seller, $16B ahead of AbbVie's Humira*. S&P Global Market Intelligence 2022 [cited 2022 September 24]; Available from: https://www.spglobal.com/marketintelligence/en/news-insights/latest-news-headlines/pfizer-s-covid-19-shot-is-2021-s-best-seller-16b-ahead-of-abbvie-s-humira-69317323.
71. Food & Drug Administration. *Drugs@FDA: FDA-Approved Drugs*. 2022 [cited 2022 September 25]; Available from: https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm.
72. Dunleavy, K. *The top 20 pharma companies by 2021 revenue*. 2022 [cited 2023 April 20]; Available from: https://www.fiercepharma.com/special-reports/top-20-pharma-companies-2021-revenue.
73. Rockhold, F., et al., *Open science: The open clinical trials data journey*. Clinical Trials, 2019. **16**(5): p. 539-546.
74. Buntz, B. *The 50 best-selling pharmaceuticals of 2022: COVID-19 vaccines poised to take a step back*. 2023 [cited 2024 January 24]; Available from: https://www.drugdiscoverytrends.com/50-of-2022s-best-selling-pharmaceuticals/.
75. Blasimme, A., et al., *Data Sharing For Precision Medicine: Policy Lessons And Future Directions.* Health Aff (Millwood), 2018. **37**(5): p. 702-709.
76. TransCelerate. *Clinical Data Transparency - De-Identification & Anonymization*. 2023 [cited 2023 January 20]; Available from: https://www.transceleratebiopharmainc.com/assets/clinical-data-transparency/.
77. ClinicalTrials.gov. *FDAAA 801 and the Final Rule (Section 801 of the Food and Drug Administration Amendments Act of 2007)*. 2022 [cited 2023 January 9]; Available from: https://clinicaltrials.gov/ct2/manage-recs/fdaaa.
78. Department of Health and Human Services. *U.S. Department of Health and Human Services. Clinical Trials Registration and Results Information Submission - A Rule by the Health and Human Services Department*. 2016 [cited 2023 January 30]; Available from: https://www.federalregister.gov/documents/2016/09/21/2016-22129/clinical-trials-registration-and-results-information-submission.
79. European Medicines Agency. *Clinical Trials Information System (CTIS): online modular training programme*. 2023 [cited 2023 January 2]; Available from: https://www.ema.europa.eu/en/human-regulatory/research-development/clinical-trials/clinical-trials-information-system-ctis-online-modular-training-programme.

80. European Union. *Regulation (EU) No 536/2014 of the European Parliament and of the Council of 16 April 2014 on clinical trials on medicinal products for human use, and repealing Directive 2001/20/EC Text with EEA relevance*. 2014 [cited 2023 January 22]; Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32014R0536.
81. European Union. *Regulation (EU) No 536/2014 of the European Parliament and of the Council of 16 April 2014 on clinical trials on medicinal products for human use, and repealing Directive 2001/20/EC Text with EEA relevance*. 2022 [cited 2022 November 4]; Available from: http://data.europa.eu/eli/reg/2014/536/oj.
82. National Institute of Health. *NIH Office of Data Science Strategy Announces New Initiative to Improve Access to NIH-funded Data*. 2022 [cited 2023 February 20]; Available from: https://datascience.nih.gov/news/nih-office-of-data-science-strategy-announces-new-initiative-to-improve-data-access.
83. White House Office of Science and Technology Policy. *Ensuring Free, Immediate, and Equitable Access to Federally Funded Research*. 2022 [cited 2023 January 4]; Available from: https://www.whitehouse.gov/ostp/news-updates/2022/08/25/ostp-issues-guidance-to-make-federally-funded-research-freely-available-without-delay/.
84. Wilkinson, M.D., et al., *The FAIR Guiding Principles for scientific data management and stewardship.* Scientific Data, 2016. **3**(1): p. 160018.
85. European Medicines Agency. *Background to clinical data publication policy* 2023 [cited 2023 January 24]; Available from: https://www.ema.europa.eu/en/human-regulatory/marketing-authorisation/clinical-data-publication/background-clinical-data-publication-policy.
86. Hemmings, R. and A. Koch, *Commentary on: Subgroup analysis and interpretation for phase 3 confirmatory trials: White Paper of the EFSPI/PSI working group on subgroup analysis by Dane, Spencer, Rosenkranz, Lipkovich, and Parke.* Pharmaceutical Statistics, 2019. **18**(2): p. 140-144.
87. Statisticians in the pharmaceutical industry. *Data Sharing SIG*. 2023 [cited 2023 July 20]; Available from: https://psiweb.org/sigs-special-interest-groups/data-sharing-working-group.
88. Bamford, S., et al., *Sharing Anonymized and Functionally Effective (SAFE) Data Standard for Safely Sharing Rich Clinical Trial Data.* Applied Clinical Trials, 2022. **31**(7/8).
89. Hoffmann, T., et al., *Focus on sharing individual patient data distracts from other ways of improving trial transparency.* BMJ, 2017. **357**: p. j2782.
90. Odame, E., et al., *Establishing a Basis for Secondary Use Standards for Clinical Trial.* Applied Clinical Trials, 2023.
91. Gomes, D.G.E., et al., *Why don't we share data and code? Perceived barriers and benefits to public archiving practices.* Proceedings of the Royal Society B: Biological Sciences, 2022. **289**(1987): p. 20221113.
92. Tedersoo, L., et al., *Data sharing practices and data availability upon request differ across scientific disciplines.* Scientific Data, 2021. **8**(1): p. 192.
93. Doshi, P. and T. Jefferson, *Clinical study reports of randomised controlled trials: an exploratory review of previously confidential industry reports.* BMJ Open, 2013. **3**(2).
94. Doshi, P., T. Jefferson, and C. Del Mar, *The imperative to share clinical study reports: recommendations from the Tamiflu experience.* PLoS Med, 2012. **9**(4): p. e1001201.
95. Hodkinson, A., et al., *The use of clinical study reports to enhance the quality of systematic reviews: a survey of systematic review authors.* Systematic Reviews, 2018. **7**(1): p. 117.
96. Wieseler, B., et al., *Completeness of Reporting of Patient-Relevant Clinical Trial Outcomes: Comparison of Unpublished Clinical Study Reports with Publicly Available Data.* PLOS Medicine, 2013. **10**(10): p. e1001526.
97. Egilman, A.C., J.S. Ross, and M. Herder, *Optimizing the data available via Health Canada's clinical information portal.* Cmaj, 2021. **193**(33): p. E1305-e1306.
98. ClinicalTrials.gov. *Submit Studies to ClinicalTrials.gov PRS*. 2023 [cited 2023 January 2]; Available from: https://clinicaltrials.gov/ct2/manage-recs.
99. Marquardsen, M., M. Ogden, and P.C. Gøtzsche, *Redactions in protocols for drug trials: what industry sponsors concealed.* J R Soc Med, 2018. **111**(4): p. 136-141.

100. Minssen, T., N. Rajam, and M. Bogers, *Clinical trial data transparency and GDPR compliance: Implications for data sharing and open innovation.* Science and Public Policy, 2020. **47**(5): p. 616-626.
101. Campbell, D., et al., *Access to unpublished protocols and statistical analysis plans of randomised trials.* Trials, 2022. **23**(1): p. 674.
102. Grady, D. and K. Thomas. *Moderna and Pfizer Reveal Secret Blueprints for Coronavirus Vaccine Trials*. 2020 [cited 2023 January 8]; Available from: https://www.nytimes.com/2020/09/17/health/covid-moderna-vaccine.html.
103. Statham, E.E., et al., *Primed to comply: Individual participant data sharing statements on ClinicalTrials.gov.* PLOS ONE, 2020. **15**(2): p. e0226143.
104. European Union. *Clinical Trials Expert Group of the European Commission representing Ethics Committees and National Competent Authorities: Good Lay Summary Practice*. 2021 [cited 2023 January 9]; Available from: https://health.ec.europa.eu/latest-updates/good-lay-summary-practice-guidance-2021-10-04_en.
105. Kirkpatrick, E., et al., *Understanding Plain English summaries. A comparison of two approaches to improve the quality of Plain English summaries in research reports.* Research Involvement and Engagement, 2017. **3**(1): p. 17.
106. Barnes, A. and S. Patrick, *Lay Summaries of Clinical Study Results: An Overview.* Pharmaceutical Medicine, 2019. **33**(4): p. 261-268.
107. EFPIA. *Implementing "Good Lay Summary Practice"*. 2022 [cited 2022 November]; Available from: https://www.efpia.eu/news-events/the-efpia-view/efpia-news/implementing-good-lay-summary-practice/.
108. World Medical Association, *World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects.* JAMA, 2013. **310**(20): p. 2191-4.
109. ClinicalTrials.gov. *Why Should I Register and Submit Results?* 2023 [cited 2023 January 15]; Available from: https://clinicaltrials.gov/ct2/manage-recs/background.
110. Turner, E.H., et al., *Selective Publication of Antidepressant Trials and Its Influence on Apparent Efficacy.* New England Journal of Medicine, 2008. **358**(3): p. 252-260.
111. Sender, D., J. Clark, and T.C. Hoffmann, *Analysis of articles directly related to Randomised trials finds poor protocol availability and inconsistent linking of articles.* J Clin Epidemiol, 2020. **124**: p. 69-74.
112. Burns, N.S. and P.W. Miller, *Learning What We Didn't Know — The SPRINT Data Analysis Challenge.* New England Journal of Medicine, 2017. **376**(23): p. 2205-2207.
113. Christian, O., et al., *Status, use and impact of sharing individual participant data from clinical trials: a scoping review.* BMJ Open, 2021. **11**(8): p. e049228.
114. Eichler, H.-G. and F. Sweeney, *The evolution of clinical trials: Can we address the challenges of the future?* Clinical Trials, 2018. **15**(1_suppl): p. 27-32.
115. Vivli. *Public Disclosures*. 2023 [cited 2023 July 20]; Available from: https://vivli.org/resources/public-disclosures/.
116. European Medicines Agency. *EU Clinical Trials Register*. 2023 [cited 2023 July 4]; Available from: https://www.clinicaltrialsregister.eu/ctr-search/search.
117. National Library of Medicine. *ClinicalTrials.gov*. 2023 [cited 2023 July 4]; Available from: https://clinicaltrials.gov/.
118. World Health Organization. *International Clinical Trials Registry Platform (ICTRP)*. 2023 [cited 2023 July 4]; Available from: https://www.who.int/clinical-trials-registry-platform.
119. Keerie, C., et al., *Data sharing in clinical trials – practical guidance on anonymising trial datasets.* Trials, 2018. **19**(1): p. 25.
120. Kochhar, S., et al., *Clinical trial data sharing: here's the challenge.* BMJ Open, 2019. **9**(8): p. e032334.
121. Rockhold, F.W., *Statistical controversies in clinical research: data access and sharing - can we be more transparent about clinical research? Let's do what's right for patients.* Annals of Oncology, 2017. **28**(8): p. 1734-1737.

122. Hosford, S.R. and T.W. Miller, *Clinical potential of novel therapeutic targets in breast cancer: CDK4/6, Src, JAK/STAT, PARP, HDAC, and PI3K/AKT/mTOR pathways.* Pharmacogenomics and personalised medicine, 2014. **7**: p. 203-215.
123. Zhang, M.H., et al., *Estrogen receptor-positive breast cancer molecular signatures and therapeutic potentials (Review).* Biomedical reports, 2014. **2**(1): p. 41-52.
124. Corona, S.P. and D. Generali, *Abemaciclib: a CDK4/6 inhibitor for the treatment of HR+/HER2- advanced breast cancer.* Drug Des Devel Ther, 2018. **12**: p. 321-330.
125. Cardoso, F., et al., *5th ESO-ESMO international consensus guidelines for advanced breast cancer (ABC 5).* Ann Oncol, 2020. **31**(12): p. 1623-1649.
126. Cancer Institute NSW. *Breast Metastatic Abemaciclib*. 2020 [cited 2021 January 18]; Available from: https://www.eviq.org.au/medical-oncology/breast/metastatic/3625-breast-metastatic-abemaciclib.
127. Ettl, J., *Management of Adverse Events Due to Cyclin-Dependent Kinase 4/6 Inhibitors.* Breast Care, 2019. **14**(2): p. 86-92.
128. Sledge, G.W., Jr, et al., *The Effect of Abemaciclib Plus Fulvestrant on Overall Survival in Hormone Receptor–Positive, ERBB2-Negative Breast Cancer That Progressed on Endocrine Therapy—MONARCH 2: A Randomised Clinical Trial.* JAMA Oncology, 2020. **6**(1): p. 116-124.
129. Dickler, M.N., et al., *MONARCH 1, A Phase II Study of Abemaciclib, a CDK4 and CDK6 Inhibitor, as a Single Agent, in Patients with Refractory HR(+)/HER2(-) Metastatic Breast Cancer.* Clinical cancer research : an official journal of the American Association for Cancer Research, 2017. **23**(17): p. 5218-5224.
130. Sledge, G.W., et al., *MONARCH 2: Abemaciclib in Combination With Fulvestrant in Women With HR+/HER2− Advanced Breast Cancer Who Had Progressed While Receiving Endocrine Therapy.* Journal of Clinical Oncology, 2017. **35**(25): p. 2875-2884.
131. Goetz, M.P., et al., *MONARCH 3: Abemaciclib As Initial Therapy for Advanced Breast Cancer.* Journal of Clinical Oncology, 2017. **35**(32): p. 3638-3646.
132. Eli Lilly and Company. *VERZENIO (Abemaciclib) - Highlights of Prescribing Information*. 2018 [cited 2021 January 11]; Available from: https://www.accessdata.fda.gov/drugsatfda_docs/label/2018/208855s000lbl.pdf.
133. Gebbia, V., et al., *Abemaciclib: safety and effectiveness of a unique cyclin-dependent kinase inhibitor.* Expert Opinion on Drug Safety, 2020. **19**(8): p. 945-954.
134. Vogenberg, F.R., *Predictive and prognostic models: implications for healthcare decision-making in a modern recession.* American health & drug benefits, 2009. **2**(6): p. 218-222.
135. Collins, G.S., et al., *Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement.* BMC Medicine, 2015. **13**(1): p. 1.
136. Johnston, S., et al., *MONARCH 3 final PFS: a Randomised study of abemaciclib as initial therapy for advanced breast cancer.* npj Breast Cancer, 2019. **5**(1): p. 5.
137. Higgins, M.I. and V.A. Master, *Who really knows the performance status: The physician or the patient?* Cancer, 2021. **127**(3): p. 339-341.
138. Bischl, B., et al., *mlr: machine learning in R.* 2016. **17**(1 %J J. Mach. Learn. Res.): p. 5938–5942.
139. Rugo, H.S., et al., *Management of Abemaciclib-Associated Adverse Events in Patients with Hormone Receptor-Positive, Human Epidermal Growth Factor Receptor 2-Negative Advanced Breast Cancer: Safety Analysis of MONARCH 2 and MONARCH 3.* The Oncologist, 2021. **26**(1): p. e53-e65.
140. Clarke, R.T., et al., *Neutropenic sepsis: management and complications.* Clinical medicine (London, England), 2013. **13**(2): p. 185-187.
141. Laskey, R.A., et al., *Predictors of severe and febrile neutropenia during primary chemotherapy for ovarian cancer.* Gynecologic Oncology, 2012. **125**(3): p. 625-630.
142. Grann, V.R., et al., *Ethnic neutropenia among women of European, African, and Caribbean backgrounds.* Journal of Clinical Oncology, 2007. **25**(18_suppl): p. 6587-6587.

143. Fontanella, C., et al., *Management of Breast Cancer Patients with Chemotherapy-Induced Neutropenia or Febrile Neutropenia.* Breast Care, 2014. **9**(4): p. 239-245.
144. Keum, J., et al., *Single-center risk factor analysis for FOLFIRINOX associated febrile neutropenia in patients with pancreatic cancer.* Cancer Chemother Pharmacol, 2020. **85**(4): p. 651-659.
145. Lyman, G.H., E. Abella, and R. Pettengell, *Risk factors for febrile neutropenia among patients with cancer receiving chemotherapy: A systematic review.* Critical Reviews in Oncology/Hematology, 2014. **90**(3): p. 190-199.
146. Chigutsa, E., et al., *Development and Application of a Mechanistic Population Modeling Approach to Describe Abemaciclib Pharmacokinetics.* CPT: Pharmacometrics & Systems Pharmacology, 2020. **9**(9): p. 523-533.
147. Lee, K.W.C., et al., *The impact of ethnicity on efficacy and toxicity of cyclin D kinase 4/6 inhibitors in advanced breast cancer: a meta-analysis.* Breast Cancer Research and Treatment, 2019. **174**(1): p. 271-278.
148. Martin, M.Y., et al., *What do cancer patients worry about when making decisions about treatment? Variation across racial/ethnic groups.* Supportive care in cancer : official journal of the Multinational Association of Supportive Care in Cancer, 2014. **22**(1): p. 233-244.
149. Shohdy, K.S., et al., *Gastrointestinal adverse effects of cyclin-dependent kinase 4 and 6 inhibitors in breast cancer patients: a systematic review and meta-analysis.* Therapeutic advances in drug safety, 2017. **8**(11): p. 337-347.
150. Thill, M. and M. Schmidt, *Management of adverse events during cyclin-dependent kinase 4/6 (CDK4/6) inhibitor-based treatment in breast cancer.* Therapeutic advances in medical oncology, 2018. **10**: p. 1758835918793326-1758835918793326.
151. Bossi, P., et al., *Diarrhoea in adult cancer patients: ESMO Clinical Practice Guidelines.* Ann Oncol, 2018. **29**(Suppl 4): p. iv126-iv142.
152. Tan, J.L., et al., *Age-Related Changes in Hepatic Function: An Update on Implications for Drug Therapy.* Drugs & Aging, 2015. **32**(12): p. 999-1008.
153. Palumbo, A., G. Lau, and M. Saraceni, *Abemaciclib: The Newest CDK4/6 Inhibitor for the Treatment of Breast Cancer.* Annals of Pharmacotherapy, 2018. **53**(2): p. 178-185.
154. Hubbard, R.E., et al., *Polypharmacy among inpatients aged 70 years or older in Australia.* Medical Journal of Australia, 2015. **202**(7): p. 373-377.
155. Faraoni, D. and S.T. Schaefer, *Randomised controlled trials vs. observational studies: why not just live together?* BMC anesthesiology, 2016. **16**(1): p. 102-102.
156. Phillips, R., et al., *Analysis and reporting of adverse events in randomised controlled trials: a review.* BMJ Open, 2019. **9**(2): p. e024537.
157. Modi, N.D., et al., *A literature review of treatment-specific clinical prediction models in patients with breast cancer.* Critical Reviews in Oncology/Hematology, 2020. **148**: p. 102908.
158. Hurvitz, S.A., et al., *Potent Cell-Cycle Inhibition and Upregulation of Immune Response with Abemaciclib and Anastrozole in neoMONARCH, Phase II Neoadjuvant Study in HR(+)/HER2(-) Breast Cancer.* Clinical cancer research : an official journal of the American Association for Cancer Research, 2020. **26**(3): p. 566-580.
159. Cesca, M.G., et al., *HER2-positive advanced breast cancer treatment in 2020.* Cancer Treatment Reviews, 2020. **88**: p. 102033.
160. Martínez-Sáez, O. and A. Prat, *Current and Future Management of HER2-Positive Metastatic Breast Cancer.* JCO Oncology Practice, 2021: p. OP.21.00172.
161. Piccart-Gebhart, M.J., et al., *Trastuzumab after Adjuvant Chemotherapy in HER2-Positive Breast Cancer.* New England Journal of Medicine, 2005. **353**(16): p. 1659-1672.
162. Swain, S.M., et al., *Pertuzumab, trastuzumab, and docetaxel for HER2-positive metastatic breast cancer (CLEOPATRA study): overall survival results from a randomised, double-blind, placebo-controlled, phase 3 study.* Lancet Oncol, 2013. **14**(6): p. 461-71.
163. Verma, S., et al., *Trastuzumab Emtansine for HER2-Positive Advanced Breast Cancer.* New England Journal of Medicine, 2012. **367**(19): p. 1783-1791.

164. Loibl, S. and L. Gianni, *HER2-positive breast cancer.* The Lancet, 2017. **389**(10087): p. 2415-2429.
165. Wang, J. and B. Xu, *Targeted therapeutic options and future perspectives for HER2-positive breast cancer.* Signal Transduction and Targeted Therapy, 2019. **4**(1): p. 34.
166. Elwyn, G., et al., *Shared decision making: a model for clinical practice.* Journal of general internal medicine, 2012. **27**(10): p. 1361-1367.
167. Black, N., *Patient reported outcome measures could help transform healthcare.* BMJ : British Medical Journal, 2013. **346**: p. f167.
168. Weldring, T. and S.M.S. Smith, *Patient-Reported Outcomes (PROs) and Patient-Reported Outcome Measures (PROMs).* Health services insights, 2013. **6**: p. 61-68.
169. Rivera, S.C., et al., *The impact of patient-reported outcome (PRO) data from clinical trials: a systematic review and critical analysis.* Health and Quality of Life Outcomes, 2019. **17**(1): p. 156.
170. Cortés, J., et al., *Health-related quality-of-life assessment in CLEOPATRA, a phase III study combining pertuzumab with trastuzumab and docetaxel in metastatic breast cancer.* Annals of Oncology, 2013. **24**(10): p. 2630-2635.
171. Welslau, M., et al., *Patient-reported outcomes from EMILIA, a Randomised phase 3 study of trastuzumab emtansine (T-DM1) versus capecitabine and lapatinib in human epidermal growth factor receptor 2–positive locally advanced or metastatic breast cancer.* Cancer, 2014. **120**(5): p. 642-651.
172. Mercieca-Bebber, R., et al., *The importance of patient-reported outcomes in clinical trials and strategies for future optimization.* Patient related outcome measures, 2018. **9**: p. 353-367.
173. Mierzynska, J., et al., *Prognostic value of patient-reported outcomes from international randomised clinical trials on cancer: a systematic review.* The Lancet Oncology, 2019. **20**(12): p. e685-e698.
174. Hopkins, A.M., et al., *Patient-reported outcomes as a prognostic marker of survival in patients with advanced nonsmall cell lung cancer treated with immunotherapy.* International Journal of Cancer, 2020. **147**(11): p. 3085-3089.
175. Quinten, C., et al., *Baseline quality of life as a prognostic indicator of survival: a meta-analysis of individual patient data from EORTC clinical trials.* The Lancet Oncology, 2009. **10**(9): p. 865-871.
176. Gotay, C.C., et al., *The prognostic significance of patient-reported outcomes in cancer clinical trials.* Journal of clinical oncology : official journal of the American Society of Clinical Oncology, 2008. **26**(8): p. 1355-1363.
177. Kerrigan, K., et al., *Prognostic Significance of Patient-Reported Outcomes in Cancer.* JCO Oncology Practice, 2020. **16**(4): p. e313-e323.
178. Badaoui, S., et al., *Patient-Reported Outcomes Predict Progression-Free Survival of Patients with Advanced Breast Cancer Treated with Abemaciclib.* The Oncologist, 2021. **26**(7): p. 562-568.
179. Tan, E., et al., *Association between Patient-Reported Outcomes and Survival in Patients with Advanced Urothelial Carcinoma Treated with Atezolizumab.* Bladder Cancer, 2021. **Preprint**: p. 1-8.
180. Badaoui, S., et al., *The predictive utility of patient-reported outcomes and performance status for survival in metastatic lung cancer patients treated with chemoimmunotherapy.* 2022, 2022.
181. Basch, E., A. Bennett, and M.C. Pietanza, *Use of Patient-Reported Outcomes to Improve the Predictive Accuracy of Clinician-Reported Adverse Events.* JNCI: Journal of the National Cancer Institute, 2011. **103**(24): p. 1808-1810.
182. Swain, S.M., et al., *Pertuzumab, Trastuzumab, and Docetaxel in HER2-Positive Metastatic Breast Cancer.* N Engl J Med, 2015. **372**(8): p. 724-734.
183. Baselga, J., et al., *Pertuzumab plus Trastuzumab plus Docetaxel for Metastatic Breast Cancer.* N Engl J Med, 2012. **366**(2): p. 109-119.

184. Diéras, V., et al., *Trastuzumab emtansine versus capecitabine plus lapatinib in patients with previously treated HER2-positive advanced breast cancer (EMILIA): a descriptive analysis of final overall survival results from a randomised, open-label, phase 3 trial.* The Lancet. Oncology, 2017. **18**(6): p. 732-742.
185. Perez, E.A., et al., *Trastuzumab Emtansine With or Without Pertuzumab Versus Trastuzumab Plus Taxane for Human Epidermal Growth Factor Receptor 2–Positive, Advanced Breast Cancer: Primary Results From the Phase III MARIANNE Study.* Journal of Clinical Oncology, 2016. **35**(2): p. 141-148.
186. Perez, E.A., et al., *Phase III, Randomised study of first-line trastuzumab emtansine (T-DM1) ± pertuzumab (P) vs. trastuzumab + taxane (HT) treatment of HER2-positive MBC: Final overall survival (OS) and safety from MARIANNE.* Journal of Clinical Oncology, 2017. **35**(15_suppl): p. 1003-1003.
187. Brady, M.J., et al., *Reliability and validity of the Functional Assessment of Cancer Therapy-Breast quality-of-life instrument.* Journal of Clinical Oncology, 1997. **15**(3): p. 974-986.
188. Smith, S.M., et al., *Clinical Cancer Advances 2021: ASCO's Report on Progress Against Cancer.* Journal of Clinical Oncology, 2021. **39**(10): p. 1165-1184.
189. ESMO. *ESMO Mission*. 2021 [cited 2021 January 15]; Available from: https://www.esmo.org/about-esmo/esmo-mission.
190. Seow, H., et al., *Development and Validation of a Prognostic Survival Model With Patient-Reported Outcomes for Patients With Cancer.* JAMA Network Open, 2020. **3**(4): p. e201768-e201768.
191. Kotronoulas, G., et al., *What Is the Value of the Routine Use of Patient-Reported Outcome Measures Toward Improvement of Patient Outcomes, Processes of Care, and Health Service Outcomes in Cancer Care? A Systematic Review of Controlled Trials.* Journal of Clinical Oncology, 2014. **32**(14): p. 1480-1501.
192. Basch, E., et al., *Overall Survival Results of a Trial Assessing Patient-Reported Outcomes for Symptom Monitoring During Routine Cancer Treatment.* JAMA, 2017. **318**(2): p. 197-198.
193. Degenhardt, T., et al., *Documentation patterns and impact on observed side effects of the CANKADO ehealth application: An exploratory analysis of the PreCycle trial.* Journal of Clinical Oncology, 2020. **38**(15_suppl): p. 2083-2083.
194. Calvo, E., et al., *PRO ASSESSMENT IN CANCER TRIALS.* Oncology PRO, 2016.
195. FDA, *Core Patient-Reported Outcomes in Cancer Clinical Trials* 2021, Department of Health and Human Services: USA.
196. Jang, R.W., et al., *Simple Prognostic Model for Patients With Advanced Cancer Based on Performance Status.* Journal of Oncology Practice, 2014. **10**(5): p. e335-e341.
197. Kelly, C.M. and A. Shahrokni, *Moving beyond Karnofsky and ECOG Performance Status Assessments with New Technologies.* Journal of oncology, 2016. **2016**: p. 6186543-6186543.
198. Bomhof-Roordink, H., et al., *Key components of shared decision making models: a systematic review.* BMJ Open, 2019. **9**(12): p. e031763.
199. Goldacre, B., *How to get all trials reported: audit, better data, and individual accountability.* PLoS Med, 2015. **12**(4): p. e1001821.
200. Gupta, Y.K., M. Meenu, and P. Mohan, *The Tamiflu fiasco and lessons learnt.* Indian J Pharmacol, 2015. **47**(1): p. 11-6.
201. Cohen, D., *Complications: tracking down the data on oseltamivir.* BMJ, 2009. **339**: p. b5387.
202. Modi, N.D., et al., *Clinical Study Report and Individual Participant Data Transparency for US Food and Drug Administration–Approved Anticancer Drugs: A Call for Systematic Data Availability.* Journal of Clinical Oncology, 2024. **0**(0): p. JCO.24.00539.

CHAPTER 10: APPENDICES

Appendix 1

Standard Request Email

Title: Data Availability and Policy Month, Date, 2021

To whom it may concern,

This is Natansh Modi, a PhD candidate in the Precision Medicine Group at Flinders University, Australia. Our group led by Dr Ash Hopkins is reviewing the availability and policy for sharing of oncology trials for independent scientific research.

I write to request clarification regarding the availability of individual participant data from clinical trials conducted by [COMPANY NAME]. I understand that individual participant data from eligible clinical trials may be shared with qualified researchers to facilitate further scientific research and honour the contribution of trial participants. Specifically, I request confirmation of whether the following clinical trials are eligible (in scope) for sharing of individual participant-level data with a valid research proposal.

1.

2.

If any of these studies are not currently in scope for sharing, I would appreciate details on when the trial(s) will become eligible for sharing individual participant data.

Thank you for your assistance with this query. The responses to this enquiry will be documented and compared to the transparency and policy of other companies. A major facet of honouring the contribution of trial participants is enabling independent research. Only through understanding eligibility for sharing across companies for these specific trials can our plans for future research be validly and timely designed.

Sincerely, Natansh Modi and Dr Ash Hopkins

Appendix 2

IPD sharing eligibility of non-industry sponsored trials

The results of 14 non-industry-sponsored trials were summarised in the product labels of the audited anticancer medicines which had been approved by the FDA in the 10-year sampling period. These trials were sponsored by 9 different non-industry sponsors. Of the 14 non-industry-sponsored trials, 2 trials were for cytotoxic medicines, 4 for immunomodulators, and 8 for targeted therapeutics not elsewhere specified. Of the 14 non-industry trials audited, 6 (43%) were in patients with solid tumours and 8 (57%) in haematological cancer. There were 7 (50%) phase 2 trials, 6 (43%) phase 3 trials, and for 1 (7%) trial the phase was not documented on clinicaltrials.gov. 10 (71%) trials had a trial start date before 1 Jan 2014, and 4 (29%) trials had a trial start date after 1 Jan 2014.

Of the 14 non-industry-sponsored trials audited, 7 (50%) trials were indicated as eligible for IPD sharing, and 7 (50%) were identified as not available for IPD sharing with independent researchers.

Appendix 3

Standardised inquiry email

To whom it may concern,

My name is Natansh Modi, and I am a researcher at Flinders University, Australia. My research focuses on data transparency policy, processes, and implications. I write to request clarification regarding the availability of Individual Participant Data from clinical trials conducted for [COMPANY NAME] products.

Specifically, I request confirmation of whether the following clinical trials are eligible (in scope) for sharing of individual participant level data with a valid research proposal. If any of these studies are not currently in scope for sharing, I would appreciate details on WHY the trial(s) are not eligible for sharing and WHEN the trial(s) will become eligible for sharing.

1. XXXXXXXX
2. XXXXXXXX

Thank you for your assistance with this query. The responses to this enquiry will be documented and compared to the transparency policies of other companies. A major facet of honouring the contribution of trial participants is enabling independent research. Only through understanding eligibility for sharing across companies can our plans for future research be validly and timely designed.

Sincerely,

Natansh Modi