Exploring the Application of a Multicenter Study Design in the Preclinical Phase of Translational Research

Victoria Hunniford

Thesis submitted to the University of Ottawa
in partial Fulfillment of the requirements for the
Master of Science in Health Systems

Telfer School of Management
University of Ottawa

© Victoria Hunniford, Ottawa, Canada, 2019

Abstract

Multicenter preclinical studies have been suggested as a method to improve potential clinical translation of preclinical work by testing reproducibility and generalizability of findings. In these studies, multiple independent laboratories collaboratively conduct a research experiment using a shared protocol. The use of a multicenter design in preclinical experimentation is a recent approach and only a handful of these studies have been published. In this thesis, I aimed to provide insight into preclinical multicenter studies by 1) systematically synthesizing all published preclinical multicenter studies; and 2) exploring the experiences of, barriers and enablers to, and the extent of collaboration within preclinical multicenter studies.

In Part One, I conducted a systematic review of preclinical multicenter studies. The database searches identified 3150 citations and 13 studies met inclusion criteria. The multicenter design was applied across a diverse range of diseases including stroke, heart attack, and traumatic brain injury. The median number of centers was 4 (range 2-6) and the median sample size was 133 (range 23-384). Most studies had lower risk of bias and higher completeness of reporting than typically seen in single-centered studies. Only five of the thirteen studies produced results consistent with previous single-center studies, highlighting a central concern of preclinical research: irreproducibility and poor generalizability of findings from single laboratories.

In Part Two, I performed semi-structured interviews with researchers who have been involved in a preclinical multicenter study. Braun and Clarkes' thematic analysis was used to identify emerging themes, and the extent of collaboration was evaluated using an established theory of collaboration developed by Wood and Gray. Twelve researchers from 6 studies were interviewed. Most participants indicated that funding and the culture of the scientific community were barriers, and that established relationships and transparency with collaborators were enablers to multicenter studies. Some participants felt that a harmonized protocol was optimal, while others stated that variability in the protocol across sites was more appropriate. Most participants indicated that multicenter studies had a purpose and place in preclinical research.

My findings also suggest that multicenter preclinical studies may provide a method to robustly assess therapies prior to considering clinical translation. These insights will allow for more effective planning and execution of future preclinical multicenter projects and may support the development of best practices and guidelines.

Acknowledgements

Collaboration is a fundamental part of my thesis not only as it was a large focus of my research, but also because the work presented in this document was only achievable through a dedicated collaborative effort. This collaboration consisted of myself and my supervisors – Dr. Agnes Grudniewicz and Dr. Manoj Lalu, to whom I would like to express my deepest gratitude. Their mentorship, support, and motivation made the work presented in this thesis possible (more realistically: I would have been completely lost without them!). With their enthusiasm and immense knowledge, I could not have imagined two better people to guide me through this Master's degree.

I would also like to acknowledge all the past and current members of the Blueprint Translational Research Team for their indispensable help and feedback, and for being a fun group of people to work with. I would especially like to thank Dr. Dean Fergusson for his expertise and valuable input on my thesis.

I am also grateful to my friends, for their genuine interest in what I've been doing over the past two years and their encouragement along the way. Whether they knew it or not, they provided me with the much-needed distraction from my research woes. After the countless explanations to their "what's your thesis about, again?" I likely developed a better understanding of my research myself (though I suspect they still don't know what my thesis is about).

Last but by no means least; I wish to express my profound gratitude to my mom and dad. I would like to thank my parents for their unfailing support and continuous encouragement – not only in my academic career, but in all interests I choose to pursue.

Table of Contents

Chapter 1: Introduction

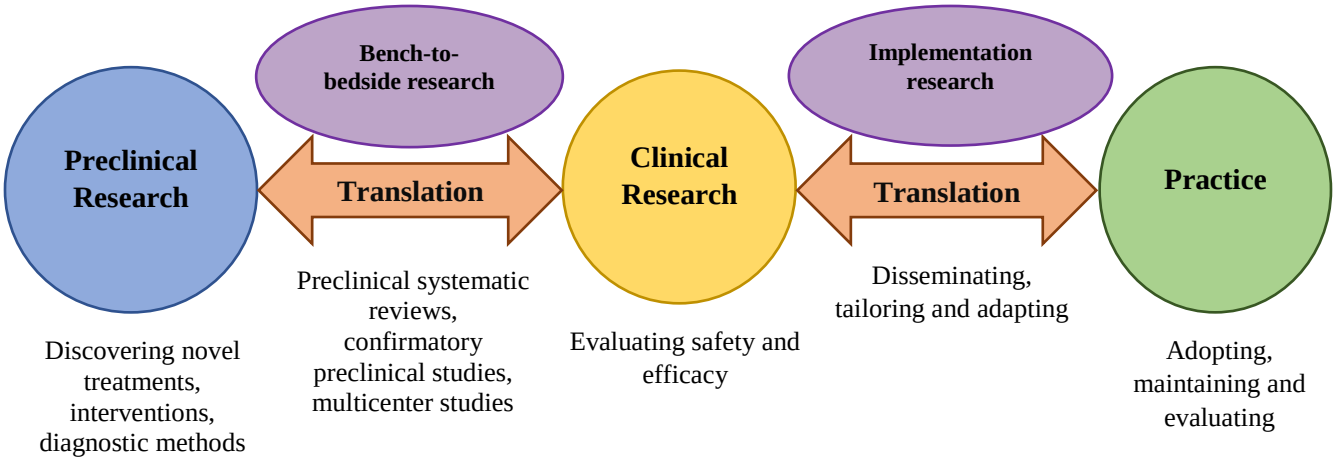
1.1 Background and rationale

Globally, billions of dollars are spent each year on biomedical, clinical and health services research [1]. These large expenses are due to the high value that society places on medical advancements, as well as the high costs to conduct research. For example, the cost of research and development of a new drug is estimated to be between $US 1 billion to $US 2.6 billion [2, 3]. The process of translating basic biomedical research (i.e. "wet-lab/bench research") to clinical practice and application (i.e. "to the bedside") involves complex steps that require tremendous amounts of time and resources [4].

Translational research – known as translational medicine or "bench-to-bedside" research, facilitates this process by providing strategies, techniques and guidelines [5]. Translational research is a field of science that incorporates both basic and applied research, and therefore involves the collaboration of basic investigators, clinicians, industry and policy makers [6, 7] and is broken down into multiple stages of research. Included in these first stages of research is the translation of basic biomedical research into clinical science, i.e., 'bench-to-bedside', which attempts to translate preclinical research at the laboratory bench to clinical trials at the patient's bedside. Preclinical research involves medically relevant investigations that are conducted using nonhuman models (e.g. cells, animals) prior to being tested in human subjects and is often the stage that precedes clinical trials. Clinical trials are a phase of translational research that tests the safety and efficacy of novel medical treatments of human diseases in human volunteers. The research following clinical trials includes the translation of trial findings into everyday practice,

and ultimately improved health. A visual depiction of a simplified version of the translational process can be seen in Figure 1.

Figure 1. – Simplified depiction of the translational research process



There has been an increasing focus on translational research in the last two decades [8]. Multiple clinical and health research studies have shown that there is a failure to translate research breakthroughs into practice and policy through clinical trials [9]. When translating a preclinical finding to a clinical therapeutic, the failure rates exceed 95% and the time-lags are significant (an average of 17 years) [10, 11]. One of the problems contributing to the translation failure of novel therapeutics may lie outside of clinical research itself, and instead originate in the preclinical stage of research [12, 13]. The issues associated with preclinical research that are believed to contribute to this translation barrier include: a) poor quality in the study design; b) poor reporting making it difficult to reproduce; c) biased selection of animal models and small sample sizes which reduces the inferential strength; d) and strong publication bias which distorts evidence to proceed to clinical trials [14-17]. In order for the process of translation to be improved, it is critical

that these issues in preclinical research be addressed so as to enhance its utility in the discovery of new medical therapies.

Many experts in the field have proposed a variety of methods to enhance the translational process and address the above-mentioned issues with preclinical research. One such method is the use of multicenter preclinical studies [18], which can be defined as collaborative research formally conducted at multiple centers with a shared protocol and analysis [14]. Although this is a relatively new concept in preclinical research, the concept of multicenter studies has been accepted for decades in clinical research. In both preclinical and clinical research, it is believed that the shortcomings of single-centered clinical trials can be improved with the use of multicenter research. Multicenter studies may increase generalizability by having greater inherent variability of samples and laboratory settings across centers. They also increase trial sample size enrolling experimental subjects across multiple centers – helping to ensure that studies are more adequately powered and providing more precise effect sizes. Additionally, these types of studies may employ extensive quality control, routine oversight, and rigorous, standardized procedures across centers, resulting in increased replicability of the study [19]. Essentially, preclinical multicenter studies may maximize transparency, improve reproducibility and rigor, enhance internal and external validity, and potentially increase the efficiency of bench-to-bedside translation [20, 21].

Several recent, high profile examples of multicenter preclinical studies have been published [20, 22-25]. Llovera and colleagues published a study in 2015 [22], where they performed a preclinical randomized controlled multicenter study to test the potential of anti-CD49d antibodies as a treatment for acute stroke. These antibodies had demonstrated great promise as a form of therapy in single-center studies by inhibiting the migration of leukocytes into the brain following acute

stroke. Six independent European research centers used harmonized, methodologically rigorous study protocols to test the antibody using two mouse models of stroke. Interestingly the antibody had no effects in one model of stroke, and only a small reduction of stroke in the other. Results from this study informed decisions to abandon further development of this therapy in its current form. This study demonstrated the potential for multicenter studies to evaluate a promising therapeutic in a rigorous, low risk of bias manner, and help determine whether further investments should be made to translate the therapy clinically.

In multicenter studies, like any project with multiple stakeholders, collaboration is a concept that can heavily impact success [26]. Collaboration occurs when two or more autonomous stakeholders work together to achieve a common goal. By sharing ideas, skills and resources, collaboration can help stakeholders maximize efficiency, reduce transaction costs and understand the research question and context being addressed more fully and in a transformative way [27]. Collaboration is particularly important in multicenter studies, as these projects not only bring together multiple independent centers, but they also assemble individuals with diverse expertise from each participating center. Therefore, multicenter studies engage in interpersonal as well as inter-organizational collaboration. As interest in multicenter preclinical trials continues to grow, a better understanding of studies that have been conducted to date, and the collaborative process involved, may help in the planning, design and conduct of such trials in the future.

1.2 Objectives

The overarching objective of the intended study is to provide better insight into preclinical multicenter studies. This may enhance the future application of multicenter study design to preclinical research. To do this, I have first synthesized the existing knowledge of these studies through a systematic review of the literature. Following this, I explored first-hand experiences

with preclinical multicenter studies through primary data collection. Specifically, I explored aspects of preclinical multicenter studies by conducting an interview study with individuals that have been involved in such studies.

Part One of my thesis determined commonalities and differences between existing preclinical multicenter studies. I produced a narrative synthesis to clarify and define study methods, elements of organizational structure, and outcomes in this field of literature. This objective has been achieved through a systematic review of published preclinical multicenter studies of controlled, interventional, *in vivo* design.

The information obtained from literature reviews is limited to what is reported in the published studies. Additionally, it is not common practice for preclinical research studies to report on or discuss experiences in preforming the study. As such, it was expected and determined that there was limited discussion on barriers, enablers and experiences with preclinical multicenter studies within the reports identified in the systematic review. Therefore, in Part Two of my thesis I performed an interview study to allow for a more in-depth investigation of preclinical multicenter studies. The goal of this interview study was to better define and explore the concepts and features of preclinical multicenter studies. The objectives that underlie the interview study and guided my thesis were:

1. Uncover the experiences and perceptions of individuals that have experience with preclinical multicenter studies.

2. From the perspectives of multicenter preclinical trial stakeholders, identify any real and perceived barriers and enablers to preclinical multicenter studies and multicenter collaboration.

3. Evaluate the extent of collaboration in preclinical multicenter studies.

Additionally, as part of the interview study, we compared and contrasted scientists' experiences and perceptions of collaboration in multicenter versus single-center studies. These objectives were achieved through in-depth, one-on-one interviews with preclinical scientists that have participated in preclinical multicenter studies.

1.3 Theory and conceptual framework

Preclinical multicenter studies require independent basic scientists to work together in a highly collaborative manner. This counters norm of preclinical research, which has traditionally been conducted in independent labs. Therefore, the concept of preclinical multicenter collaboration is an unexplored domain. For this reason, I first evaluated the concept of collaboration in the literature to better understand how it has been studied previously. Through this informal search of the literature it was determined that collaboration has been studied extensively in many fields of research. Additionally, it appears that the concept of collaboration has not been explored in basic science or preclinical research.

Theory of collaboration

Collaboration is a complex concept embedded in social norms across various social units and disciplines. As such, collaboration merits theoretical exploration - building upon previous work that has been done in numerous fields such as business management and social sciences [28, 29]. Therefore, I used a theory of collaboration developed by Wood and Gray [27] to help develop and support the design of the interview study. This theory is based upon theoretical contributions from case research and preexisting theories in various domains of resource dependence, corporate social performance/institutional economics, strategic management/social ecology, microeconomics, institutional/negotiated order, and political theories. It has been applied

extensively in business research, social sciences and political research [30-32]. The theory states that "collaboration occurs when a group of autonomous stakeholders of a problem domain engage in an interactive process, using shared norms, rules, and structures, to act or decide on issues related to that domain". There are six individual elements in collaboration: *the stakeholders*; *stakeholder autonomy*; *the interactive process; shared norms, rules, and structures*; *the action*; and *domain orientation*. *The stakeholder*s refer to the group or organizations with an interest in the problem domain. The theory states that the stakeholders must remain autonomous throughout the course of the collaborative process, in that they must retain their independent decision-making powers, though they may agree to relinquish some autonomy. If participants relinquish all autonomy, the authors state that the organizational form created is not collaboration, but instead a merger. *The interactive process* indicates that a change-oriented relationship exists, and that all stakeholders are involved in this relationship. *Shared norms, rules, and structures* of collaboration must also be agreed upon explicitly by stakeholders and can continuously evolve over the course of the collaborative venture. *The action* implies that participants must intend to 'act or decide', in that the goal of the collaboration is directed towards an objective. Lastly, *domain orientation* requires that the collaborators orient their processes, decisions, and actions towards issues related to the problem domain that brought them together [27].

This theory provides a formal definition of collaboration but does not aim to make assumptions about which or how many stakeholders will participate, at what level of social organization the collaboration will occur, whether the collaborative structure will be temporary, the nature of the intended outcome, or whether the effort will succeed. To evaluate the extent of collaboration in preclinical multicenter studies, I assessed real examples of preclinical multicenter groups to determine if all six elements are present in these collaborative alliances [27, 33].

Features of collaboration

Wood and Gray further expand their theory from the previously discussed definition of collaboration, to include additional features that may or may not be present in a collaboration alliance. This includes a convener, as well as potential precursors to, and outcomes of collaboration

It is stated that the presence of a 'convener' is another condition that facilitates the formation of an alliance, though it is not necessary. Conveners identify and bring all the legitimate stakeholders to the table. A convener may possess informal authority such as that based on position and influence in an informal network, expertise and knowledge with respect to the problem domain, or credibility among the stakeholders of the domain [27].

The theory also addresses the question of why collaboration is initiated. In regard to environmental complexity and control, the authors suggest that organizations collaborate to reduce and control environmental uncertainty and turbulence, yet they acknowledge that collaborative alliance could very well increase the complexity and possibly the uncertainty and turbulence of an organization's environment [27, 33]. It is hypothesized that high stakes and high interdependencies are two necessary factors that motivate parties to collaborate. This is based upon Resource Dependence theory [34], in that interdependences are created because organizations possess or control vital resources, and thus are the source of environmental pressure for one another. Organizations seek to reduce these pressures and manage the interdependencies by gaining control over crucial resource supplies, thus reducing uncertainty of gaining those supplies. Collaboration can help organizations achieve this objective of improving the efficiency of resource use [27]. It was unknown if these elements apply to multicenter studies, therefore by

exploring collaboration in these studies I also evaluated the potential applicability of this theory to such collaborative situations.

Another hypothesized motivation for initiating collaborative alliances is the need for individual organizations to maximize efficiency and reduce transaction costs. Furthermore, partners in a collaboration have the potential to understand the problem domain more fully and in a transformative way: in collaborative alliances, the process of building a joint appreciation enables all stakeholders to increase their understanding of the problem by learning the desired and intended actions of others. It is also acknowledged that collaboration may increase transaction costs for organizations, introduce them to bilateral and multilateral relationships to which they must attend, require them to develop new skills and abandon or reshape old ones. These outcomes of collaboration are likely case dependent and would need to be explored in greater depth to determine if they hold true. Therefore, this thesis also examined the possible outcomes or consequences of collaboration in multicenter studies.

As mentioned previously, collaboration is an essential component of multicenter studies. In both single-center and multicenter studies, the research center(s) (i.e. laboratory) at which the project is being conducted is made up of many individuals with various backgrounds and expertise, and specific roles. These individuals include principal investigators, research associates, technicians, students, and those with administrative roles. Each possess a different set of expertise essential to a research project and must collaborate in order to execute the study with success. Common across single-centered and multicenter research projects, is interpersonal and intra-organizational collaboration. Multicenter studies add greater complexity to the collaborative component, as these studies also add inter-organizational collaboration. Therefore, it is important that the theory of collaboration used in this thesis address these three different levels of collaboration.

Competing theories of collaboration

Numerous theories of collaboration exist, but after critically considering several of these theories it was determined the Wood and Gray's theory of collaboration is most suited to studying collaboration in preclinical multicenter studies. This theory works well for the intended study because it covers both institutional and interpersonal aspects of collaboration, and it gives a broad definition of the construct. Other theories have mainly focused on inter-organizational and intergroup collaboration, with little attention to collaboration on the interpersonal level [35, 36]. Though Wood and Gray's theory addresses interpersonal aspects of collaboration, such as the role of an influential and highly regarded individual in establishing collaboration, they do not reduce collaboration into forms of leadership, followership and teamwork – as Colbry, Hurwitz and Adair [37] do in their collaboration theory. Another factor that was essential in Wood and Gray's theory is its inherent flexibility to be applied to a range of collaborative scenarios. They devised their theory by extracting several overarching theoretical underpinnings from nine research-based articles and two overviews - all of which used collaboration in distinct ways and vastly different fields of research [30-32]. Other authors developed their theories from more specific scenarios, and thus had narrower ranges of application. For example, Patel, Pettitt and Wilson [38] developed their theory from analyzing collaboration in aerospace, construction, and automotive sectors. Therefore, their theory was very specific to collaboration in these fields, and it would be difficult to apply to preclinical multicenter studies.

Another notable work by Thomson and Perry[36], further developed Wood and Gray's theory into a framework of the process of collaboration, consisting of five key dimensions to collaboration. They also named Wood and Gray's collaborative process component of their theory, the 'Antecedent–Process–Outcome Framework'. Though this work is a more developed version of

Wood and Gray's original theory, I chose not to use this framework for my thesis as it is focused on the collaborative process in public administration and services and is applicable to collaboration between public managers. Furthermore, unlike the original theory – which has been applied to a number of different domains, it appears that the framework has not been used in areas outside of public administration. Thus, I felt that this framework was not appropriate to evaluate collaboration in preclinical multicenter studies and elected to use Wood and Gray's more general theory.Wood and Gray's theory of collaboration was used to shape and define the approach applied to address the objectives of Part Two of this thesis, mainly the third objective – the extent of collaboration. The theory aided in the development of the interview guide, questions for respondents that were aimed at specifically addressing collaboration. The first two objectives of the interview study are broader issues that do not specifically explore collaboration, and therefore did not need to apply the theory as extensively through a deductive approach. The research objectives one and two were addressed using an inductive approach.

Chapter 2: Methods

The methods section is divided into two parts, in which I describe the methods for the systematic review and the interview study, respectively.

2.1: Systematic Review

I developed the systematic review protocol in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses for protocols (PRISMA-P) [39]. The systematic review protocol for Part One of the thesis was developed in collaboration between myself, my two thesis supervisors (Dr. Agnes Grudniewicz and Dr. Manoj Lalu), and a senior scientist with expertise in knowledge synthesis (Dr. Dean Fergusson). The protocol was registered with the International Prospective Register of Systematic Reviews (PROSPERO CRD42018093986) and was posted on the Collaborative Approach to Meta-Analysis and Review of Animal Data from Experimental Studies (CAMARADES) website (http://syrf.org.uk/protocols/).

Eligibility Criteria

Population

The population of interest was preclinical, interventional, multicenter, and controlled or comparison studies. *Preclinical* was defined as research that is medically relevant and is conducted using nonhuman models prior to being tested in human subjects. *Multicenter* was defined as cooperative research formally conducted between multiple research centers (sites). Models were limited to *in vivo* experiments but were not limited by the clinical scope or domain of the preclinical study. Furthermore, selected studies were not limited by the individual studies' type of intervention, protocol, variables of interest or results. Observational, *in vitro*, *ex vivo*,

clinical studies and veterinary trials were excluded. The rationale for including interventional, *in vivo* studies was that they are more likely aimed at translating therapies from animal models into clinical trials.

Intervention, comparators, outcomes

There were no limitations to specific intervention, comparator or outcomes of individual studies included.

Study design

Eligible preclinical studies included *in vivo*, controlled, interventional studies of randomized and non-randomized designs. *In vivo* experiments needed to be conducted at two or more independent sites for the study to qualify as multicentric. The sites needed to also share more than just general study objectives to be considered multicentered. Features that met the multicenter criteria included, but were not limited to: shared design, protocol, the animal model, the method of analysis, and the primary endpoints tested with or without identical measurement apparatuses; separate centers for coordination, protocol development, and data analysis; and study objective, timelines, protocols, and dissemination strategies developed *a priori*. Veterinary clinical trials, *in vitro* and *ex vivo* studies (with no *in vivo* component), and retrospective data analysis from multiple sites were excluded.

Search strategy

I developed the search strategy in collaboration with an information specialist (Risa Shorr, The Ottawa Hospital) and my thesis supervisors. Electronic databases Embase (Embase Classic and

Embase 1947 – 30 January 2018), and MEDLINE (Ovid MEDLINE Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid Medline Daily and Ovid Medline 1946 to 29 January 2018) were searched on January 31, 2018. The search strategy was developed using keywords related to multicenter studies as well as preclinical research, such as preclinical OR animal model experimental model, AND multicenter OR cross laboratory. The search strategy was tested through the inclusion of six target articles [17, 20, 22-25] identified by my thesis supervisor (Dr. Manoj Lalu) prior to the systematic search. A second, independent librarian peer reviewed the search strategy according to the Peer Review of Electronic Search Strategy (PRESS) framework [40]. After all relevant studies were identified from the first search, additional studies were found in the references of the included studies. Therefore, another search was run on April 24th, 2018 using search terms that aimed to include the missed studies. A final search was run on May 8th, 2019, as there was another missed study identified during an interview with a scientist who was involved with one of the previously identified preclinical multicenter studies. For all three searches, no study scope, date or language limits were imposed. The most up-to-date search strategy is presented in the supporting information (Appendix 1), as well as the PRESS review (Appendix 2).

Screening and data extraction

The results from the literature search were uploaded to Distiller Systematic Review Software (DistillerSR®; Evidence Partners, Ottawa, Canada). DistillerSR is a cloud-based program that facilitates the review process by managing studies through customized screening, auditing and reporting. Duplicate references were removed and two reviewers – another Masters student (Emma Grigor) and I, independently screened titles and abstracts based on the eligibility criteria.

Any disagreements were resolved by consensus. For the second stage of screening, my supervisor and I independently screened the full text reports of included references based on the eligibility criteria. Disagreements were resolved via consensus.

Data was extracted using a standardized extraction form developed in DistillerSR that was piloted by myself on five studies. The extraction form was revised based on feedback from my thesis supervisors and was then made into an Excel spreadsheet where extracted data was saved. Data included characteristics of the studies: publication details (authors, year published, journal), the country(ies) where the study was conducted, sources of funding, the number of centers involved (experimental and non-experimental), the disease model, animal species and sex, sample size, treatment/exposure, primary and secondary study outcomes, the reported results, statements of barriers and facilitators, and statements of recommendations and suggestions for future testing of the specific therapy being investigated.

Assessing completeness of reporting and risk of bias

Risk of bias and quality of reporting were assessed independently by my thesis supervisor and I, and disagreements were resolved via consensus. All randomized, interventional studies were assessed as high, low, or unclear for the 10 domains of bias from the SYRCLE "Risk of Bias" assessment tool for preclinical *in vivo* studies [41]. The 'other sources' of risk of bias domain was divided into 4 sub-domains (funding influences, conflicts of interest, contamination, and unit of analysis errors). An overall 'other' risk of bias assessment was given based on the following: overall high risk of bias if one or more of the four other sources were assessed as high; overall unclear risk of bias if two or more of the four other sources were assessed as unclear; and overall

low risk of bias if three of the four other sources were assessed as low, with the fourth assessed as unclear.

Quality of reporting of the studies was assessed using a checklist modified from various sources: consolidated Standards of Reporting Trials (CONSORT) [42]; the National Institutes of Health (NIH)'s principles and guideline for reporting preclinical research [13]; and the Good Clinical Practice (GCP) Guidance Document: E6(R2) [43]. The checklist is provided in the supporting information (Appendix 3) with details on the sources for each item.

Assessing degree of collaboration

An assessment of the degree of collaboration reported was made for each multicenter study. The degree of collaboration assessment criteria I developed was based on a validated scale used to measure collaboration among grant partners. The original scale developed by Frey and colleagues used five levels of collaboration ranging from 1 to 5, and a score of 0 for no interaction at all [44]. Assessments were given based on the reporting of (or lack thereof) study design elements in 3 domains: protocol development, protocol execution, and collaboration. These elements have been suggested to contribute to success in multicenter clinical, and early preclinical experiments, such as the use of separate centers for data processing and coordination (coordination domain), a collaborative development and shared protocol between all centers involved (protocol development domain), standardized reporting across all centers involved (protocol execution domain) [9, 14, 15, 26]. The degree of collaboration in each domain was assessed with a four-point scale (0-3) defined as low, medium, high, or unclear and scored from 1 (for low) to 3 (for high). Domains that were unclear were given a score of 0. The summary table used

color to visualize the degree of collaboration across the three domains (red = low, yellow = medium, green = high, white = unclear) (Appendix 4). Degree of collaboration was assessed independently by myself and my supervisor.

Results and data synthesis

The results from the systematic review were extracted descriptively and data was interpreted through a narrative synthesis. Descriptive data was synthesized and presented through tabulation of textual elements [45]. Studies were arranged in tables to report on study design and basic characteristics, and risk of bias assessments. A synthesis of any statements and examples pertaining to barriers and facilitators in conducting a multicenter study was also performed on the included multicenter research studies. Given that the focus of this review is reporting design and characteristics rather than a specific outcome (and the heterogeneous nature of preclinical study topics) performing a meta-analysis was not appropriate. All results were reported in accordance with the PRISMA guidelines [46]. A copy of the PRISMA checklist is provided in Appendix 5.

Deviations from protocol

The original protocol submitted to PROSPERO indicated that the degree of collaboration would be evaluated through a summed score of all three domains, resulting in a numerical value ranging from 0 to 9. After expert feedback, it was decided not to use a quantitative score and to assess the domains individually. We elected to assign colors for each assessment to provide a visual representation of the degree of collaboration across the three domains.

2.2: Interview Study

In Part Two of this thesis, I conducted an interview study in order to answer the three objectives of the thesis through an in-depth, qualitative approach. As such, the questions that guided the interview study were:

- What are the experiences, and perspectives of individuals that have experience with preclinical multicenter studies?
- What are the real and perceived barriers and enablers to preclinical multicenter studies and multicenter collaboration?
- What is the extent of collaboration in preclinical multicenter studies?

Ethics

Ethics approval was obtained from the Ottawa Hospital Research Ethics Board (20180768-01H) and the University of Ottawa (S-11-18-1388). Written informed consent was obtained from interview participants and all procedures followed institutional guidelines.

Participants and recruitment

Purposeful and snowball sampling methods were used to recruit participants for this study. The intended sample of participants were preclinical scientists who have had experience with (i.e. conducted or participated in) a preclinical multicenter study. This sample was identified in the systematic review (Part One) of published preclinical multicenter studies. The corresponding authors of the identified 13 papers were contacted through email and were asked to participate in an interview. These papers came from eight distinct groups, where one group published five individual papers, one group published two papers, and the remaining six papers were published by six independent groups. Those who were interested in participating in the study responded to

the recruitment email and were contacted by me to schedule an interview at a time convenient to them. Semi-structured interviews were conducted by telephone. Additional participants were recruited through snowball sampling methods, where the corresponding authors who participated in an interview were asked if they could provide the contact information of other members involved in their respective multicenter groups.

Interview guide development

An interview guide was developed prior to the start of the interview process (Appendix 6). The interview guide explained why the participants were contacted and asked them questions that were specific to the preclinical multicenter study they had conducted - including their experiences, and their views and opinions of preclinical multicenter studies in general. Other questions included those that aimed to understand participant perspectives on how single-center studies compare with multicenter studies. If a participant began to discuss their views on a different topic or question naturally, I would adjust the sequence of questions as necessary to improve the flow of the discussion. Lastly, I collected participants' demographic information.

The guide also contained questions aimed to evaluate the extent of collaboration within the preclinical multicenter study. These questions were developed using Wood and Gray's theory of collaboration [27], where the six elements of the theory were worked into questions in the interview guide to elicit whether or not these elements were present in the collaboration.

The guide contained a total of 23 open ended questions with prompts to allow for elaboration. The interview guide was refined based on feedback from a preclinical laboratory technician (Casey Lansdell) and my thesis supervisor who is a preclinical principal investigator (Dr. Manoj Lalu), to be appropriate for the preclinical study context. Four pilot interviews were completed

with experts in both the interview process and in preclinical research. There were no major revisions to the guide at this stage.

Data analysis

In order to monitor the progress of the interviews and to permit follow up of issues that may have emerged from the previously conducted interviews, an iterative process was used. As such, interviewing, transcription and analysis occurred concurrently during the interview phase of the thesis. During this process, I recorded and transcribed all interviews verbatim. Interview transcripts were imported to NVivo 11 (QSR International, Doncaster, Australia) – a standard qualitative software program designed to facilitate the coding of data.

My thesis supervisors (Drs. Agnes Grudniewicz and Manoj Lalu) and I independently coded the first two transcripts, compared coding and resolved discrepancies. A code book was developed using theory-informed deductive analysis and inductive thematic analysis[47]. The deductive analysis applied Wood and Gray's theory of collaboration, where the six elements of the theory were used as the deductive codes in the code book. The inductive codes emerged from the first two coding sessions, where all three researchers independently identified similar constructs then agreed upon their names and definitions. A third transcript was independently coded by two researchers using the code book, compared for discrepancies and the code book was finalized. Following this, I independently coded all remaining transcripts. The code book contained 18 codes; 6 deductive and 12 inductive codes. Eight of the inductive codes were further broken down to contain sub-codes, with 31 in total. The code book and sub-code book with definitions can be found in Appendices 7 and 8.

After coding, analysis of data was separated by inductive and deductive analysis. Summaries of each inductive code and sub-codes were created and sorted into a table. Several exemplary

quotes for each of the codes and sub-codes were extracted to support the summaries. From these summaries emerging themes were identified and recurring themes that cross-cut codes were grouped to create new key themes. A framework analysis[48] was also performed using the inductive codes to understand and present the different structures of the preclinical multicenter studies. Each quote was followed by a participant ID, in the form of a number and letter. The number corresponds to the participant (numbered 1 to 12) and the letter corresponds to the study group they belonged to (lettered A to F).

The deductive analysis used Wood and Gray's theory of collaboration to answer Objective 3 of my thesis: What is the extent of collaboration in preclinical multicenter studies?

To determine the extent of collaboration I used the theory's definition of collaboration, containing six separate elements: *stakeholder autonomy*; *interactive process*; *shared norms, rules, and structures*; *the action*; *domain orientation*; and *a convener*. These elements were the six deductive codes used during the coding process. During the analysis, I assessed whether the six elements were present within each multicenter group. If data that aligned with a deductive code was identified in the interview transcripts, the corresponding element of collaboration would be considered to be present. The extent of collaboration in each multicenter group was determined based on the presence of these elements, where no elements indicated that it was not truly a collaboration, and the presence of all six elements suggests a fully collaborative alliance. If the multicenter group contains some but not all elements, it would be considered a collaboration to a certain degree. Summaries of each of the deductive codes (six elements) were sorted into a table with illustrative quotes from various participants.

To further evaluate the extent of collaboration, I used the seven features of collaboration that were identified and discussed in the theory: *high stakes and interdependencies; transaction*

costs; understand the problem more fully; maximize the efficiency of resource use; introduction to bilateral and multilateral relationships; *skills – developing new and abandoning/reshaping old*; and *environmental complexity and control*. I looked at the inductive and deductive codes to determine if each feature is present within the multicenter groups and how it is manifested. High stakes and interdependencies between stakeholders, decreasing transaction costs, understanding the problem more fully, and maximizing the efficiency of resource use are four potential reasons for why stakeholders choose to collaborate. Introduction to new relationships and developing, abandoning or reshaping skills are two possible outcomes due to collaboration. Environmental complexity and control, and transaction costs may increase or decrease as a consequence of the collaboration. I evaluated whether or not these features were present in the data, using a framework analysis looking at the seven features across the six multicenter groups. I determined whether a feature was present if participants expressed them as reasons for collaboration (four reasons), if they indicated it was an outcome due to the collaboration (two direct outcomes), or if there was an increase or decrease in the feature as a result of collaborating in the preclinical multicenter studies (two feature that may increase or decrease).

Chapter 3: Results

The results section is divided into two parts in which I report the results of the systematic review and the interview study individually.

3.1: Systematic Review

For Part One of my thesis, I performed a systematic review of all published preclinical multicenter studies. Here I report the results of the systematic review and discuss how this data contributed to answering the three research objectives of my thesis.

Search results and study characteristics

The database searches identified a total of 3150 papers after duplicates were removed (Figure 2). Two additional papers were identified through a search of references of included papers. Thirteen articles met eligibility criteria following title, abstract, and full-text screening (Tables 1 and 2).

The identified studies fell into six clinical domains: traumatic brain injury (n = 5), myocardial infarction (n = 2), stroke (n = 2), traumatic injury (n = 2), diabetes (n = 1), and effects of stimulant exposure (n = 1). Nine of thirteen studies were published in 2015 and 2016; two studies were published in 2009, and two in 1985, 1999. Eleven studies were fully funded by government sources; two studies had partial funding from government and charitable or academic funding. Three studies were international (studies with centers located in the USA, Germany, France, Canada, Finland, Hungary, Italy, the United Kingdom, and Spain), and ten studies were conducted solely in the USA (all centers located in the USA). The median number of total centers involved per multicenter study was 4 (range: 2-6), and the median number of experimental centers performing *in vivo* work was 3 (range: 2-5). Nine studies (69%) reported

having non-experimental centers involved, such as a coordinating center, data processing center, biomarker core, and a pathology core. Five different species of animals were used by the studies: mice (*n* = 5), rats (*n* = 5), swine (*n* = 3), rabbits (*n* = 1), and dogs (*n* = 1). One study used three species of animals for their experiments. The median sample size was 133 (range 23-384 animals), and a total 1854 animals were used across the thirteen studies, 90% of which were lab rodents (mice and rats).

Figure 2. - Preferred reporting items for systematic reviews and meta-analysis (PRISMA) flow diagram for study selection.

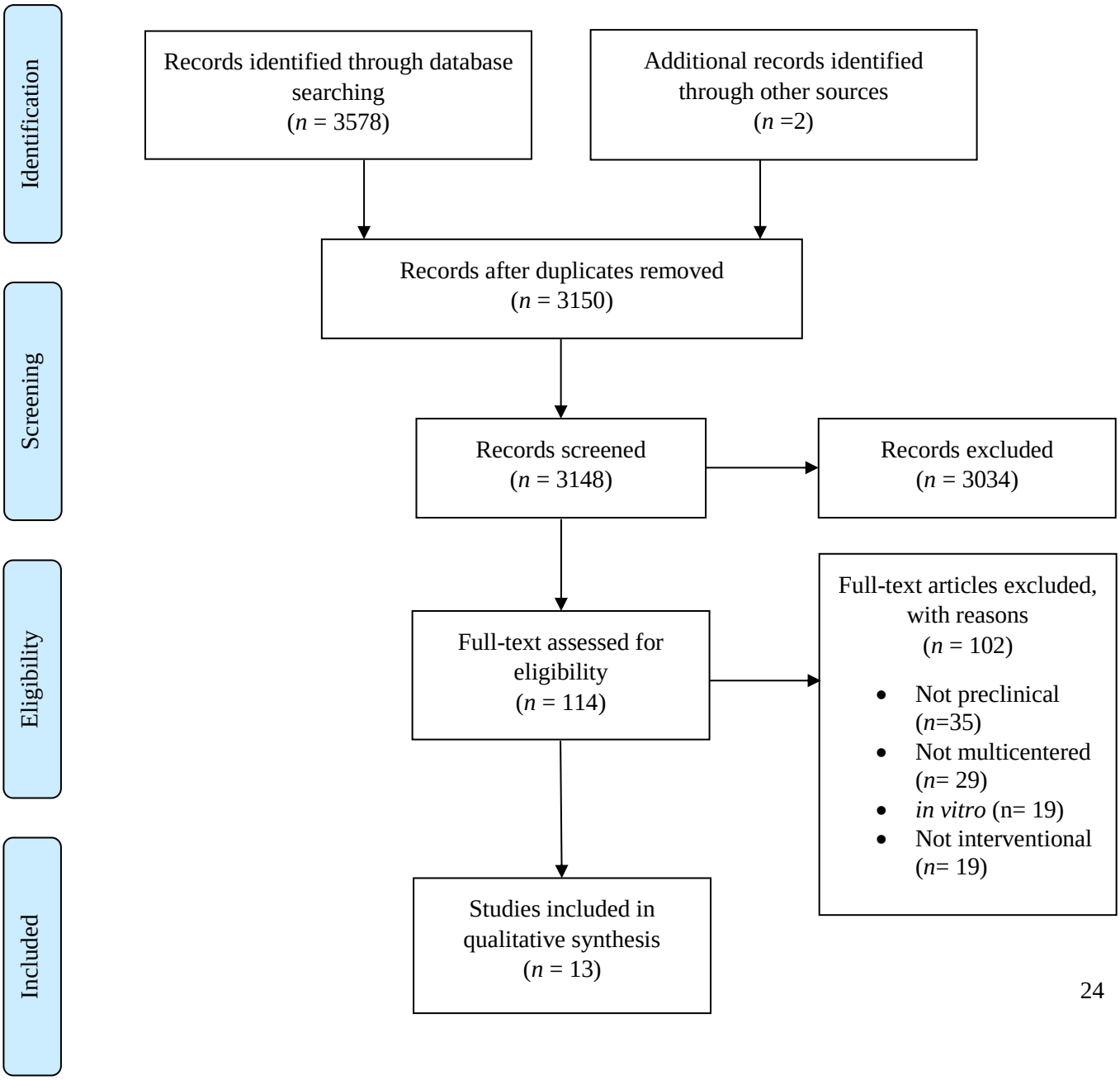


Table 1. - Basic study characteristics of preclinical multicenter studies

Author, Year	Center location	Journal	Funding	Centers Performing *In Vivo* Work	Non-experimental centers*	Animal, Sex	Sample size
Reimer, 1985	US	*Circulation Research*	Government (NHLBI)	3	1	Dog, both	51
Crabbe, 1999	Canada, US	*Science*	Government	3	0	Mouse, both	384
Alam, 2009	US	*Journal of Trauma: Injury, Infection and Critical Care*	Government (US army)	3	0	Swine, F	60
Spoerke, 2009	US	*Archives of Surgery*	Government (US army)	2	0	Swine, NA	32
Jones, 2015	US	*Circulation Research*	Government (NHLBI)	3	3	Mouse, M Rabbit, M Swine, F	47 23 26
Llovera, 2015	France, Germany, Italy, Spain	*Science Translational Medicine*	Government, academic and charitable	5	1	Mouse, M	315
Maysami, 2015	Finland, France, Germany, Hungary, UK, Spain	*Journal of Cerebral Blood Flow & Metabolism*	Government (FP7/ 2007-2013, INSERM), academic	5	1+	Mouse, M	241
Bramlett, 2016	US	*Journal of Neurotrauma*	Government (US army)	3	1	Rat, M	140
Browning, 2016	US	*Journal of Neurotrauma*	Government (US army)	3	1	Rat, M	130
Dixon, 2016	US	*Journal of Neurotrauma*	Government (US army)	3	1	Rat, M	135
Gill, 2016	US	*Diabetes*	Government, charitable	4	0	Mouse, F	NA
Mountney, 2016	US	*Journal of Neurotrauma*	Government (US army)	3	1	Rat, M	128
Shear, 2016	US	*Journal of Neurotrauma*	Government (US army)	3	1	Rat, M	142

Legend: FP7/2007-2013 – European Union Commission 7th Funding Programme; INSERM - Institut national de la santé et de la recherche médicale; NHLBI – National Heart, Lung, and Blood Institute; UK – United Kingdom; US – United States

***Non-experimental center: A site/lab not involved with the in vivo experiment (data processing, coordinating, biomarker, or pathology centers)**

+Center that was both an experimental center and a coordinating center

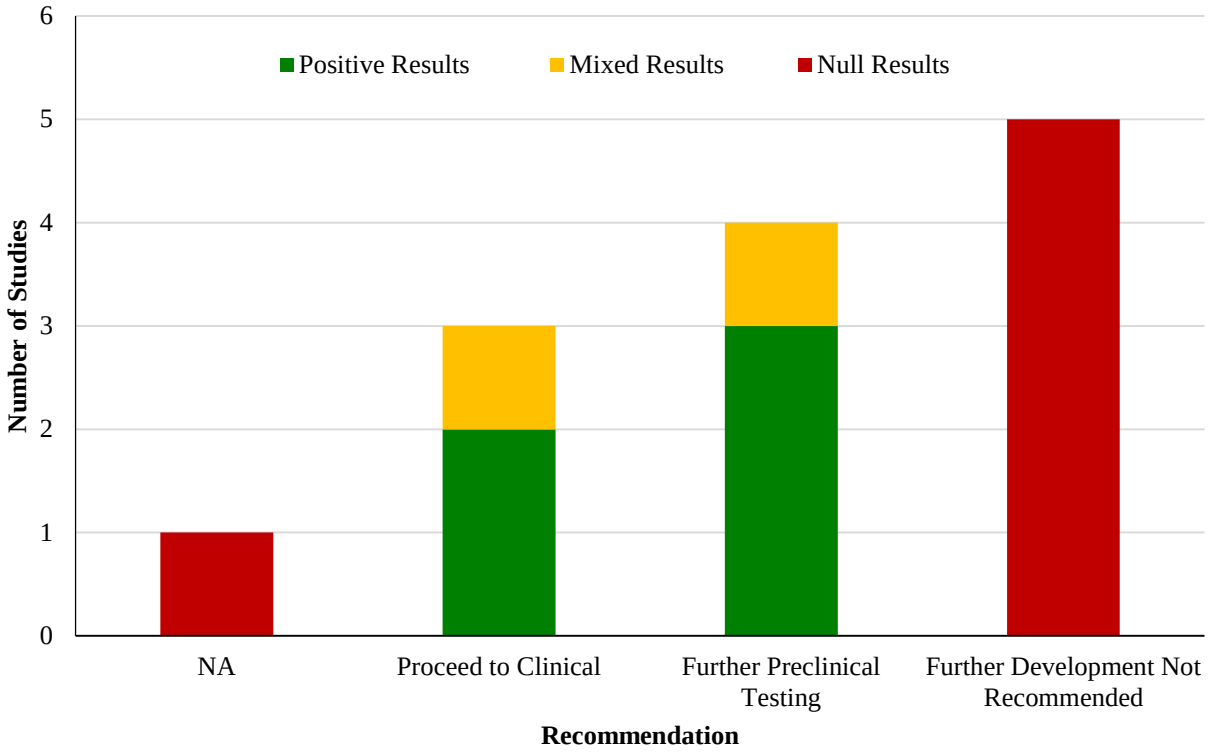
Table 2. - Study design characteristics of preclinical multicenter studies

Author, Year	Disease model	Treatment	Primary Outcome	Secondary Outcomes	Reported Results	Recommendations for Future Research
Reimer, 1985	Myocardial infarction	Verapamil and ibuprofen	Infarct size	Mortality, hemodynamic measures, pathological/histological features, regional blood flow	Null	Not reported
Crabbe, 1999	Stimulant exposer	Cocaine	Locomotor activity		Mixed	Further preclinical testing
Alam, 2009	Polytrauma	Blood transfusion	Hemodynamic parameters	Mortality	Positive	Further preclinical testing
Spoerke, 2009	Traumatic injury/hemoragic shock	Lyophilized plasma	Residual clotting activity	Mortality, hemodynamic measures, total blood loss, coagulation profiles, inflammatory measures	Positive	Further preclinical testing
Jones, 2015	Myocardial infarction	Ischemic preconditioning	Infarct size	Hemodynamic measures, regional blood flow, heart weight, troponin I, mean arterial pressure	Positive	Further preclinical testing
Llovera, 2015	Stroke	Anti-CD49d antibody	Infarct size	Functional outcome, invasion of leukocytes to brain	Mixed across models (positive, null)	First-in-human clinical trial
Maysami, 2015	Stroke	Interleukin-I receptor antagonist	Infarct size	Odema, functional outcome, mortality	Positive	Extensive clinical trial
Bramlett, 2016	Traumatic brain injury	Erythropoietin	Cognitive outcomes, biomarkers, motor outcomes, neuropathology		Null	No further preclinical study
Browning, 2016	Traumatic brain injury	Levetiracetam	Cognitive outcomes, biomarkers, motor outcomes, neuropathology		Positive	Further preclinical testing and first-in-human clinical trial
Dixon, 2016	Traumatic brain injury	Cyclosporine	Cognitive outcomes, biomarkers, motor outcomes, neuropathology		Null	No further preclinical testing
Gill, 2016	Diabetes	Combined anti-CD3 + IL-1 blockade	Blood glucose		Null	Pause clinical trial
Mountney, 2016	Traumatic brain injury	Simvastatin	Cognitive outcomes, biomarkers, motor outcomes, neuropathology		Null	No further preclinical study
Shear, 2016	Traumatic brain injury	Nicotinamide	Cognitive outcomes, biomarkers, motor outcomes, neuropathology		Null	No further preclinical study

Reported outcomes

Five of the studies (38%) reported that the treatment showed statistically significant, positive results; six studies reported that the treatment showed non-significant or null results; two studies reported that the results were mixed (positive and null) across different animal models of the disease of interest [22], or outcome measures [24] (Table 2). Based on their respective results, twelve studies made explicit statements of recommendations, or future directions for the therapy of study. Six studies stated that they would conduct further testing or recommended that further preclinical testing be done on the tested therapy. Four studies indicated they would not continue testing or recommended that no further preclinical testing be done. Three studies recommended proceeding with clinical trials. Four of the studies that recommended further preclinical testing had mixed (n = 2) or positive (n = 2) results; the three studies that recommended to proceed with human clinical trials had mixed (n = 1) and positive results (n = 2), and the five studies that suggested that there should be no further testing (clinical or preclinical) all had null results (Figure 3). Brief synopses of the thirteen studies can be found in supporting information (Appendix 9), along with sample statements of their future recommendations (Appendix 10).

Figure 3. - Preclinical multicenter studies reported recommendations for future studies along with the reported results of each study.



Risk of Bias

None of the 13 studies (0%) were considered low risk of bias across all ten domains (Table 3). Ten studies randomized animals to experimental groups and two of these reported the method of random sequence generation. Nine studies had a low risk of detection bias by blinding of outcome assessors. Eight studies were at low risk of performance bias by blinding personnel administering interventions. All but one study was unclear if animals were randomly housed during the experiments. Five studies from the same research consortium (Operation Brain Trauma Therapy) had high risk of bias for 'other sources' of bias due to potential industry-related influences (Table 3). The four 'other sources' of risk of bias assessments for each study can be found in the appendices (Appendix 11).

Table 3. – Risk of bias assessment of multicenter preclinical interventional studies

Study	Sequence generation*	Baseline characteristics	Allocation concealment*	Random housing	Blinding of personnel	Random outcome assessment	Blinding of outcome assessment	Incomplete outcome data*	Selective outcome reporting	Other sources of bias+*
Reimer, 1985	U[1]	U	L	U	H[3]	U	L	U	L	U
Crabbe, 1999	U[1]	L	U	L	U	U	U	L	L	H
Alam, 2009	U[1]	L	U	U	U	L	U	U	L	U
Spoerke, 2009	U[1]	L	U	U	U	L	U	U	L	U
Jones, 2015	L	L	U	U	L	L	L	L	L	L
Llovera, 2015	L	L	U	U	L	L	L	L	L	L
Maysami, 2015	H[2]	H	L	U	L	L	L	U	H	H
Bramlett, 2016	U[1]	U	U	U	L	L	L	U	L	H
Browning, 2016	U[1]	U	U	U	L	L	L	U	L	H
Dixon, 2016	U[1]	U	U	U	L	L	L	U	L	H
Gill, 2016	H	L	U	U	U	L	U	U	U	L
Mountney, 2016	U[1]	U	U	U	L	L	L	U	L	H
Shear, 2016	U[1]	U	U	U	L	L	L	U	L	H

Legend: H = High risk of bias (red), L = Low risk of bias (green), U = Unclear risk of bias (yellow)
*Items in agreement with the Cochrane Risk of Bias tool [49]
Baseline Characteristics: Low risk = Relevant baseline characteristics equal between experimental groups or controlled for. Unclear = Relevant baseline characteristics unreported. High risk = Relevant baseline characteristics unbalanced between experimental groups and not controlled.
Random Housing: Low risk = Animal cages were randomly placed within an animal room/facility, Unclear = Housing placement unreported, High risk = Animals place in non-random arrangement in animal room/facility.
Blinding of Outcome Assessment: Low risk = Outcome assessors were blinded to the study groups when assessing endpoints/animals Unclear = Insufficient information to determine if outcome assessors were blinded during assessment. High Risk = Outcome assessors not blinded to the study groups.
Incomplete Outcome Data: Low risk = N values were consistent between methods and results for the outcomes. Unclear = N value was either not presented in the methods or in the results, and therefore there is insufficient information to permit judgement. High risk = N values were not consistent between methods and results for the outcomes.
Selective Reporting: Low risk = The methods section indicated pre-specified outcome measure. Unclear: Was not clear about the pre-specified primary endpoints and outcome results. High risk = The outcome was presented in the results but not pre-specified in the methods section.
[+] other sources include funding influences, conflicts of interest, contamination, unit of analysis errors
[1]Method of randomization not specified
[2]Some centers used appropriate randomization where others used pseudo-randomization
[3]Assessed as high because one arm of the study was inadvertently unblinded

Reporting quality

Overall, completeness of reporting across all thirteen studies was high, ranging from 66% to 100% of checklist items being reported. One of the thirteen studies reported on all the 29 items in the preclinical multicenter reporting checklist. The domains with the highest completeness of reporting included replicates (biological vs. technical), statistics, blinding, and discussion (Table 4). The domains for standards, randomization, sample size estimation, and inclusion/exclusion criteria were variable in the completeness of reporting. The introduction and abstract domain had the lowest completeness of reporting, as eight of the thirteen studies did not report that the study was multicentered in the title (or use a synonym such as consortium, cross-laboratory, or multi-institutional; item 1, Table 4) and less than half indicated the number of participating centers in the abstract. Though it was not an item in the reporting checklist, four studies included *preclinical* in the title, two studies had *animal (or swine) model* in the title, and five studies from the same traumatic brain injury consortium did not indicate that the study was preclinical (or synonym) in the title. Additional papers [50, 51] that accompanied the five traumatic brain injury studies included *preclinical* in the paper title. Reporting assessment for all twenty-nine items across the thirteen studies can be found in the appendices (Appendix 12).

Table 4. – Frequency of reported preclinical multicenter checklist items

Domain	#	Item Description	% of studies that reported
Intro/ abstract	1	Identification as a multicenter study in title	38
	2	Abstract states number of participating centers	46
Standards	3	Community based reporting guidelines listed	15
	4	Names of each participating center listed	100
	5	List roles of participating centers (central coordinating center, experimental site)	85
	6	No changes, or if applicable major changes to study protocol after commencement are documented	100
Replicates (biological vs. technical)	7	Results substantiated by repetition under a range of conditions at each site	100
	8	Number of subjects per outcome	100
	9	Number of measurements per subject for one experimental outcome stated	85
	10	Number of subjects per center	77
Statistics	11	List of the total number of subjects used in each experimental group	85
	12	List of all statistical tests used	100
	13	Definition of the measure of central tendency	100
	14	Definition of the measure of dispersion	100
Randomization	15	Random group assignment reported	100
	16	Description of the method of random group assignment	31
Blinding	17	Experimenters blinded to group allocation during conduct of the experiment	69
	18	Experimenters blinded to group allocation during result assessment	77
Sample Size Estimation	19	Description of an *a priori* primary outcome	100
	20	Sample size for each site computed during study design	31
	21	Description of the method of sample size determination	31
Inclusion and Exclusion Criteria	22	Total number of animals for the experiment reported	54
	23	Description of the criteria used for the exclusion of any data or subjects	46
	24	List losses and exclusions of animals at the end of experiment	54
	25	All outcomes described, or description of any outcomes that were measured and not reported in the results section	100
	26	Previous or pilot/preliminary studies performed and listed	85
	27	Results were significant, or if not, null or negative outcomes included in the results	100
Discussion	28	Limitations of the study are documented	85
	29	Discrepancies in results across centers expected or absent, or if not, they discussed	100

Legend: Coloured cells indicate the frequency (%) of item reported over all included studies. Frequency (%) ranges: 0-37 = red; 38-76 =yellow; 77-100 = green.

Reported barriers and facilitators

Five of the thirteen studies (38%) explicitly reported on the barriers and facilitators to conducting a multicenter study. The most frequently reported barrier identified in all five studies was the establishment of a consistent protocol, with attention to exact experimental details across research centers [20, 22, 23, 25, 52]. In addition to the challenge of the initial protocol development, studies reported difficulty in centers strictly adhering to the established protocol throughout the entirety of the study. One study [23] had considerable issues in adhering to the protocol, and in effect had to modify their methods through the course of the study.

Three studies [20, 22, 23] reported that differences in equipment and resources across centers as a barrier. This made it difficult to conduct a collaborative project and to communicate what measurements and endpoints would be assessed. Specific experimental conditions that investigators were unable or unwilling to modify included animal models of the disease, animal housing conditions, the separate labs' operating and measurement procedures, equipment, and institutional regulations. There was also inconsistent funding across research centers. Different centers had separate budgets with different amounts of funding that could be allocated to the study. If the protocol was to be harmonized, then it had to be adapted to fit each center's budget accordingly (i.e., the center with the smallest budget set the spending limit) [23]. Alternatively, centers developed a general protocol but adapted it to fit their own respective budget with what resources they had. Another barrier identified was ethics approval for animal experiments at all the centers [22]. This was especially significant when centers were located in multiple countries, as each country had different regulations for ethical approval [22, 23].

Jones *et al.* [20] suggested that a clearly defined experimental method was facilitated by employing a pilot test through all the centers and subsequently developing a protocol collaboratively.

Developing a defined experimental protocol also included establishing an agreed upon timeline, laboratory setup and method of analysis and measurement. Maysami *et al.* and Reimer *et al.* [23, 25], retrospectively suggested that a similar approach might have enhanced the conduct of both of their studies. Another study reported that another facilitator was the use of a centralized core for administration and data processing [20, 22]. The validity of reports depends on the control statistical and data management; and having one center coordinate these operations reduces the chances of error or bias in the analysis. Other facilitators were related to the interpersonal aspect of collaboration. These included having investigator leadership through regular conferences and check-ins from beginning to end of the project [53] and building upon previously established personal/professional relationships between investigators [20].

Degree of collaboration

Overall, the thirteen studies scored medium to high in the degree of collaboration (Table 5). The 'development' domain of collaboration had the greatest number of studies given 'high' degree of collaboration assessment. The 'execution' and 'coordination' domains had similar overall degree of collaboration assessments, though there was one more study assessed as 'low' in the coordination domain, and one more study assessed as 'high' in the execution domain. One study [20] had high degree of collaboration across all three domains, and one study received an unclear score of 0 across all three domains [54] – meaning that this study was unclear in the reporting for each domain of collaboration.

Table 5. – Degree of collaboration assessment

Author, Year	Development	Execution	Coordination
Reimer, 1985	2	2	3
Crabbe, 1999	3	3	2
Spoerke, 2009	0	0	0
Jones, 2015	3	3	3
Llovera, 2015	2	3	3
Maysami, 2015	2	1	1
Bramlett, 2016	3	2	2
Browning, 2016	3	2	2
Dixon, 2016	3	2	2
Gill, 2016	2	3	1
Mountney, 2016	3	2	2
Shear, 2016	3	2	2

3.2: Interview Study

Here I report the results of the interviews from Part Two of my thesis and describe how this data answered the three research objectives of my thesis.

Sample characteristics

Twelve preclinical researchers participated in an interview. Interview duration ranged from 21 to 54 minutes, with a median interview length of 36 minutes. Participants conducted or participated in six of the eight independent preclinical multicenter groups identified in the systematic review. (One group published five of the multicenter studies, another group published two of the studies, and six individual groups published one multicenter study each. Collectively, this makes up all thirteen of the preclinical multicenter studies). No participating researcher could be recruited from the remaining two multicenter research groups. The corresponding author (Gill, 2016) did not respond to my recruitment and follow-up emails, and all researchers of the study published in 1985 were retired or deceased (Reimer, 1985).

Participants (eight male; four female) had a median age of 54 years (range: 35-72). Participants lived in four different countries: the USA (*n* = 6), Germany (*n* = 4), the UK and France (*n* = 1 each) and studied in six different areas of medical research: stroke (*n* = 6), traumatic brain injury (*n* = 3), MI, polytrauma, and psychology (*n* = 1 each). The descriptive statistics of the twelve participants are presented in Table 6.

Table 6. - Participant demographics (*n* = 12)

Demographic characteristic	Descriptive Statistic	*n*
Age range	35 to 72 years (median: 54)	
Gender	Male	8
	Female	4
Interview length range	21 to 54 minutes (median: 36)	
Countries of residence	United States	6
	Germany	4
	France	1
	United Kingdom	1
Field of research	Stroke	6
	Traumatic brain injury	3
	Myocardial infarction	1
	Polytrauma	1
	Psychology	1

Coding summary and data saturation

A minimally adequate sample size for reaching data saturation was estimated with guidance from the literature [55-57]. Collectively, the literature recommends a sample size between 10 to 16 participants for homogenous groups as well as to use personal judgement in determining whether

new information is emerging from the data. We considered the sample of participants a homogenous group because they are all well-established researchers working with animal models of human diseases.

Objective 1: Experiences and perceptions of multicenter studies

The first objective of this thesis was to uncover the experiences and perceptions towards preclinical multicenter studies from those who have participated in such a study. To provide context for the experiences and perceptions, I first describe the study logistics, structures and processes of the multicenter research groups; followed by the roles and responsibilities within them. Following this, I illustrate the participants' views, opinions and experiences with their respective study and of preclinical multicenter studies in general.

Study logistics, structures, and processes

All participants explained the details and logistics of their respective multicenter studies. One of the most reoccurring details discussed was that of study funding. Two of the groups did not receive funding for their studies, while two other groups received complete funding from the military. One of the groups had complete funding for the first phase of their project, but their grant was not renewed to complete the subsequent phases of the project. The sixth group received partial funding for their study, but contributed just as "much of their money, pulled from other grants" (4, B).

Another key topic that was discussed throughout many interviews was protocol harmonization between participating study sites. The discussion around protocol harmonization was centered on the shared study details across sites or details participants felt should harmonize as much as possible. This included analgesics used; method, dose and timing of drug administration; animal model and species; and the methods of outcome measurement and analysis. Several participants

believed that the protocols should be as harmonized as possible, while others thought that there should be some degree of variation. There were a few participants that elaborated further into this, articulating that the level of harmonization depends on the goal of the research question and the goal of the study. There were differences in the extent and degree of protocol harmonization across collaborating sites, which could be sorted into three categories: full harmonization – where every detail of the study protocol was shared across sites (three groups); harmonization with multiple models – where the standard operating procedures were harmonized, but different sites had different experimental models of the disease (two groups); and not harmonized – where participating laboratories were harmonized only in the intervention tested and the primary outcomes (one group).

"I would very much advocate for the point of trying to harmonize protocols very much, as much as possible across study sites to get a homogenous model and homogenous read-out. And I think that this is the main problem which needs to be tackled in future studies like this kind" (1, A)

"I would rather think it's better to have and accept some diversity in the protocols" (9, D)

Participants also discussed the structural features of the multicenter studies. As was presented in the results of the systematic review in Part One of this thesis, the six multicenter groups varied in structure by the number of participating experimental sites, the different roles at each site, and the use of non-experimental sites. One structural feature that was employed across three of the six study groups was the use of central analysis of outcomes. The participants who spoke of this feature deemed that it was a critical aspect of their respective studies' structure and discussed this feature in relation to making the preclinical multicenter studies as similar to a clinical trial as possible. Other structural aspects that were mentioned by participants included the use of non-experimental sites for analysis of biomarkers (a measurable characteristic as an indicator of the

physiological state [58]), histology (study of microscopic anatomy; analysis of tissue samples) and pathology (study of disease progression and effects; analysis of tissues, cells and body fluids) an extensively detailed manual of operations, experimenter training sessions, and technician site visits to the participating laboratory prior to the start of the study.

Communication between collaborating sites was another logistical aspect that was discussed and was deemed important by the participants. All participants stated that site principal investigators and other study personnel communicated through email, participants from four of the six groups indicated that there were regular or monthly telephone or videoconference calls between all sites, and one group met in person at their dedicated preclinical multicenter consortium meetings.

Other details and logistics of the multicenter studies that were discussed in lesser detail included: site selection and study length. Some participants mentioned basing selection of collaborating sites upon two things: the collaborators having previously established relationships, and the participating collaborators' skill and expertise in the field of research. Discussion around study length was focused on the time it took to develop the protocol, and the time from the conception of the idea to the dissemination of the results. At least one participant from all the independent groups indicated that protocol development took between 3 and 12 months. Other participants indicated that the study length from conception to dissemination was a longer process than in their own single-center study experience for various reasons such as: time to obtain a grant for the project, time to develop and harmonize the protocol, and agreeing on how to present and write up the results.

The last logistical feature that was discussed was dissemination of the results. All participants shared the number of papers published from their multicenter projects; three groups published two papers (the study results and a separate commentaries/editorial), one group published the

results and then published a paper that elaborated on the methods several years later, one group published five papers on the results with three papers on methods, discussion and commentary, and one group published only the results of their study. Three participants from two groups mention that they published their results in high impact journals and received "notoriety" for it.

The group leadership structure was also discussed. There were two types of leadership styles across the six groups: shared leadership - where no one person assumed a leadership role across or within collaborating sites; or hierarchal leadership – where there was one site or principal investigator leading and spearheading the project. Participants from three of the groups stated (two stated explicitly, one implied) that leadership was shared equally between collaborating principal investigators once the study was initiated. The three other groups appear to have a hierarchal leadership structure. A summary of the groups' logistics and structures is found in Table 7.

"I think what we did well was to share leadership across the group. So, this enabled us in the end to find a consensus on every single topic. And some of them of course were more debated heavily than others" (1, A)

"I don't think you could call any of our roles a leadership role. Like all three of us were partners in this study" (4, B)

"Any future consortium would definitely need to make sure that they have a very specific leader picked out with some very specific leadership traits. Having one specific person responsible for the overall centers is good. If it's just a generic group where there's no definitive leader… to keep everybody on track, it's very very easy for things to stop happening because there's no

specific person going 'we need to focus on what we had originally set out to do'... it has to be one person" (6, C)

Table 7. - Study logistics and structures of the six multicenter groups

Participant	Group	Funding	Protocol harmonization	Experimental sites	Non-experimental sites	Papers published	Type of leadership
1, 2, 3	A	Funded by collaborators	Fully harmonized	5	1	2 (study & elaboration)	Shared
4	B	Funded by collaborators and small grant	Fully harmonized	3	0	2 (study & elaboration)	Shared
5, 6, 10	C	Funded by the military	Multiple models	3	1	9 (5 studies & 4 elaboration)	Hierarchal
7, 8, 9	D	Funded by collaborators	Not harmonized	5	0	1 (study)	Hierarchal
11	E	Funded by the military	Fully harmonized	2/3*	0	2 (studies)	Shared
12	F	Funded by the government	Multiple models	3	3	2 (study & elaboration)	Hierarchal

* Two sites initially, a third site joined during the second phase

Roles and responsibilities

One of the findings that emerged from the interviews was that there were distinct roles that individuals maintained throughout their preclinical multicenter study. These roles were both formal roles - established prior to the start of the studies, or informal roles which developed over the course of the collaboration. In all, there were five roles that were identified across participants: *convener*, *figurehead*, *site-point person*, *non-experimental specialist*, and *experimenter*. All participants identified and discussed an individual who acted as the *convener* for their respective groups. This individual was someone who was well-connected, and who identified the collaborators/laboratories, invited the collaborators to join the multicenter group, and was essentially responsible for initiating the multicenter collaboration.

"I initiated and coordinated the whole study… I approached a few collaborators and asked them if they would be happy to join such a study" (1, A)

"And I'm very happy that he took charge and just had the pressure or the urge to put his [drug forward]. And he also convinced some others to bet on that one, so to put some money into it" (2, A)

Another role that was identified by multiple participants was the *figurehead*. This individual was someone who was well-respected and a leader and/or expert in the field of preclinical study or of the multicenter study design. In two of the groups, those which were the largest preclinical consortiums, the role of figurehead and convener was shared by the same individual.

A third role that was discussed by participants from four of the groups was that of the *site point-person*. This role was either held by one person across the entire research group or one person per site; and could be, but was not exclusively a site principal investigator. This role involved

managing study personnel and resources, being the go-to person for logistical questions and problems, and channeling communication across and within participating sites.

"I think at each site, you needed the point person that was going to basically take care of the protocol, make sure the study is running smoothly, make sure all personnel are in place, and that all the outcome measures are being done correctly. I think at each study site, there was a person that did that and I think that helped a lot with making sure that everything was getting done in terms of the drug testing" (5, C)

A fourth role that was discussed by some participants was the *non-experimental specialist*. This role was present in the groups that employed separate non-experimental sites (a lab/site not performing the preclinical experiment). The people that fell under this role included study and data coordinators, data analysists, and protocol review committees. Though only three of the groups indicated that they employed these personnel, participants from two of the studies that did not have these roles, indicated that in future preclinical multicenter studies they would enlist a study coordinator and data analysts.

"He was just the coordinator, in terms of putting things on the website, calls for proposal and that. And we had an external advisory committee also, which updated, for example if we tested some drug, and then the results were presented to the external advisory committee, the coordinator was the one who presented these results there" (12, F)

Intuitively, the fifth role – *experimenter*, was present in all the study groups. These were the individuals that actually performed the preclinical experiments and included the laboratory technicians and graduate students. The roles present within each of the six multicenter groups are shown in Table 8.

Table 8. - The five roles present in each of the six preclinical multicenter groups.

Participant	Group	Convenor	Figurehead	Site point-person	Non-experimental specialist	Experimenter
1, 2, 3	A	Yes	Yes	Yes	Yes	Yes
4	B	Yes	Yes	No	No	Yes
5, 6, 10	C	Yes	Yes	Yes	Yes	Yes
7, 8, 9	D	Yes	Yes	No	No	Yes
11	E	Yes	Yes	Yes	No	Yes
12	F	Yes	Yes	Yes	Yes	Yes

Another facet of roles that was discussed was the collaborators' personality traits. There were twelve different personality traits that were mentioned by one or more participants: open-minded, anal compulsive, independent minded, curious/inquisitive, friendly, professional, collegial, multi-tasker, willing to participate, organized, willing to reach consensus, and team player. These traits were often discussed as being enablers to collaborations or preclinical research.

"Everyone was extremely open, and no one complained that there was blinding. And everyone was eager to learn about the results" (2, A)

"We're really all about team science here with my group and we feel like this is – team science and collaboration is how we're going to push drugs forward for example, into the clinic" (5, C)

Participants also discussed responsibilities within their multicenter groups. In addition those associated with the five roles, there were three major responsibilities that were discussed: *protocol development*, *manuscript writing*, and *obtaining funding*. Similar to the differences in the level of protocol harmonization across sites, the level of shared responsibility for *protocol development* and *manuscript writing* differed throughout the multicenter groups. In five of the

groups, it appeared that these tasks were shared – with minor differences in the level of input across collaborating sites. *Protocol development* appeared to be the task most equally and thoroughly shared across collaborating sites, where participants felt that this was part of the "collaborative nature" of a multicenter study. *Manuscript writing* was shared by having the first draft of the manuscript written by one of the site principal investigators, and then was passed on to the other collaborating sites. In the case of the study group that performed multiple studies – each with its own manuscript, each site was responsible for one study manuscripts each. This level of sharing is different from that of *protocol development*, as it appeared that the protocol was drafted and finalized between all sites concurrently. The study group that did not have a harmonized protocol did not have shared responsibility in developing the protocol nor in the writing of the manuscript. In this study, one site took full responsibility for *protocol development* and *manuscript writing*.

A third additional task that was shared was *obtaining funding* – either through shared grant writing if the study was funded, or by leveraging funds from non-related grants. All three of the study groups who were self-funded appeared to have equal contributions of funds from all collaborators. Of the three groups who received funding, it was disclosed that the grant writing was led by the *convener/figurehead.* Though it is likely that the other sites within these three groups contributed to writing the grant, the extent of this contribution was not discussed. The sharing of the three tasks across the six different groups is presented in Table 9.

Table 9. - The three shared tasks across the six preclinical multicenter groups.

Participant	Group	Protocol development	Obtaining funding	Manuscript writing
1, 2, 3	A	Yes	Yes	Yes
4	B	Yes	Yes	Yes
5, 6, 10	C	Yes	Spearheaded grant writing	Yes
7, 8, 9	D	No	Yes	No
11	E	Yes	No, one site provided funding	Yes
12	F	Yes	Spearheaded grant writing	Yes

Experiences, opinions and views

Above, I reported on participants' practical experiences with study logistics, structures and processes within their respective multicenter groups. Now I discuss participants' personal experiences and their opinions and views of preclinical multicenter studies. In general, all participants had positive attitudes towards, and experience with their respective preclinical multicenter collaborations. All participants stated that it was a positive experience or that they enjoyed collaborating. Most participants indicated that they felt confident in their ability to collaborate and perform a multicenter study and that it was a worthwhile experience. They gave various reasons for why they felt it was a positive experience and discussed their original expectations and the benefits they received by participating in the studies.

Participants experienced and received different benefits from participating in their multicenter collaboration, which were grouped into three categories: *learning*, *reaching goals* and *rewards*. *Learning* occurred when a participant either learned something specific from the collaboration or generally saw the multicenter study as a learning experience. Most participants stated that they learned more about their own field, the differences in techniques across sites or about the value

and purpose of multicenter studies. All participants that identified *learning* as a benefit stated that it would help them in their future research endeavors.

"I would say that maybe the outcome isn't different, but the quality of the work has improved because we have learned so much from each other, and our models are very well refined now" (12, F)

Reaching goals occurred when participants felt that their respective group met their explicit study objectives, or any ancillary objectives they had for the multicenter group. Almost all participants indicated that *reaching goals* was achieved and was one of the benefits in their respective multicenter studies. Participants noted that they also reached goals regardless of their study results (positive or negative/null) by achieving ancillary objectives, such as; proving a preclinical multicenter study is possible, successfully mimicking a multicenter clinical trial, or replicating results across sites with a harmonized protocol. The third benefit that many participants experienced was *rewards*. These *rewards* included establishing and solidifying relationships with members of the multicenter group; getting publications; and exposure and publicity. Multiple participants stated that participating in the multicenter research group helped them advance their career and connected them with individuals outside of their usual collaborations.

"It enhanced relationships with other labs I think. It put me in touch with other labs that were doing cross-center studies and I've given a talk on this several times, it was quite useful as we got exposure. I think it made me realize and understand some of the complexities of models and some of the limitations. And also, I think it helped us refine our experimental approaches" (5, C)

All participants expected that their studies would run smoothly, and they had not anticipated any major problems. Participants held such positive expectations due to social and personal enablers. These enablers are explained in greater detail in the results of Objective 2 below. Some participants stated that they expected minor "hiccups", but nothing they could not overcome. Some did experience challenges that they did not foresee but were "only daily business problems" (1, A).

Though all participants held an overall positive attitude towards their respective studies, as well as multicenter studies in general, there were some participants that mentioned a negative experience with their respective multicenter study. These could be categorized into three different groups: *regret*, *stress* and *incidents*. Overall, most participants indicated that they had very few, inconsequential *regrets* about their multicenter collaboration: two stating that their regrets were in regard to minor study details that are experienced in many areas of research.

"The little things…those general regrets that almost every study has. Regrets that are just generic science regrets." (6, C)

Two participants from the study group with a non-harmonized protocol stated that they regretted having not fully harmonized their protocol as this caused them issues during the course of their experiment.

Several participants discussed the *stress* associated with their multicenter collaboration, though many indicated that conducting a multicenter study was not more or less stressful than a single-centered study. Aspects that participants identified as being stressful included: harmonizing every detail of the experimental protocol with agreement from each participating site, obtaining funding for a large study, and publishing the results in an appropriate journal. Overall, it did not

appear that *stress* was a major influence on attitude towards preclinical multicenter collaborations.

The last negative contributing factor identified was unexpected *incidents*. Several participants mentioned having experienced unexpected *incidents* within their multicenter collaboration, which included having variable results between centers when protocols were not fully harmonized and obtaining null results from the tested therapy when the supporting single-center research showed that the therapy was efficacious. In addition, the group using a non-harmonized protocol had unexpected difficulty with publishing the results, as participants from this group stated that the journals saw issues with their lack of harmonization. One group had unexpected backlash from the scientific community, as it was stated that demonstrating variability in laboratory conditions having an effect on results was like "airing the dirty laundry" and "undermined all validity behind behavioural mouse testing" (4, C). Another group had complications in transporting study material from one site to another when crossing European country borders and clearing customs.

Comparing multicenter studies to single-center studies

Another approach employed in the interviews encouraged participants to discuss their experiences and opinions of multicenter studies was to have them compare their experiences with their single-centered research.

Half the participants indicated that multicenter studies and single-centered studies cannot be directly compared because they are distinctly different. These participants explained that both types of studies have their own place and purpose in preclinical research and that one design is not superior to the other; as both are necessary in the translational process.

"I think they're designed to ask different questions. So I think you can't say that one is better over the other, they're too different… So I think the scientific question is a different one, so therefore you can't compare in my view. The techniques are the same and the experimental methods are the same, but the rationale behind them is different" (7, D)

The other half of the participants did make a comparison and stated that single-centered studies are cheaper and easier to initiate and perform. When asked if participants had a preference between working on either a preclinical multicenter study or single-center study, all participants indicated that they did not prefer one study design over the other. Overall, no participants indicated that the multicenter study design was superior to the single-center study design; though some felt that multicenter studies were superior for confirmatory research prior to testing a therapy in a first-in-human clinical trial.

"The job of single-centered trial is not something that can ever be achieved in a multicenter trial. A multicenter validation trial cannot identify a new target, it can just validate something" (1, A)

Regardless of the type of response a participant gave when asked to compare the two study designs, all participants indicated that individual principal investigators have more personal control in single-centered research, and less control in multicenter studies. In their single-center research, participants felt that they could do what they wanted and had more flexibility throughout the course of their experiments. In their experience with multicenter studies, they did not have this freedom as the multicenter studies were less flexible and more rigorously controlled. From these discussions, it appears that multicenter studies have more regulated control, but individual principal investigators have less personal control.

The preclinical multicenter study design

Participants also commonly discussed the purpose and appropriateness of the preclinical multicenter study design in preclinical research in general and in their respective fields of study. All indicated that the overall purpose of preclinical multicenter studies is to translate preclinical therapies into clinical trials. Many participants went on to give reasons for why multicenter studies can and should be used in the translational process. These reasons included: increased sample size, robust, rigor and quality control, greater generalizability, replication across different laboratory environments, and more expertise and knowledge with multiple collaborators. The majority of participants also stated that multicenter studies should be used for a validation or confirmatory study of a treatment or therapy, and not be used to test a new mechanism nor in exploratory or early basic preclinical research.

"The issues with translations could be due to the preclinical stage, and this could be that a lot of the time the studies are not robustly conducted and they're irreproducible. So by applying a multicenter approach similarly that's used in clinical practice, experts believe that this may increase the robustness of the studies, there will be more quality control and it could increase the reproducibility of the study, and test the generalizability of new treatments and therapies in preclinical models" (11, E)

Another way in which the participants discussed their thoughts on appropriate methods and design of multicenter studies was by comparing them to clinical trials. Several participants made a comparison to clinical trials, either stating that they modelled their multicenter study after one or that future preclinical multicenter studies should implement the methods and procedures employed in clinical trials. Specific examples that were mentioned included having

standardization and a shared protocol across sites, routine oversight and a monitoring committee or individual, external analysis of outcomes, and complete blinding and randomization.

Future research and contributions to science

All participants mentioned and discussed future research using the preclinical multicenter study design – either their own research, in their field of study, or in different areas where the multicenter design could be applied. All but one, who is retired, state that they would be involved in or would conduct another multicenter study. When asked if there was anything that they would have changed about their respective multicenter groups or the experiments themselves, most participants affirmed that there was nothing they would do differently. Some said that they would only change minor details in the experimental procedure, and one stated that they would like to have a fully harmonized protocol (this from the study group with a non-harmonized protocol).

Multiple participants also discussed their multicenter studies as a contribution to their field of preclinical research, or how the multicenter design is a contribution to translational research of therapeutics. Some stated that the multicenter design was a critical and necessary stage in the translation process.

"I think it's something that we have to do for the future. I think that all drugs that we expect to translate to clinic should be tested first in this type of multicentric preclinical trial" (1, A)

Objective 2: Barriers and enablers to conducting preclinical multicenter studies

The second objective of this thesis was to uncover and illuminate the barriers and enablers to initiating and conducting a preclinical multicenter study. This objective was touched upon in the systematic review in Part One of the thesis, but limited insight into these aspects can be learned

from the published reports of these studies as it is not standard practice for studies to report on the barriers and enablers experienced in the study.

Barriers

All 12 participants mentioned and discussed barriers they experienced in their respective multicenter studies or groups, and/or barriers to multicenter studies in general. These were sorted into three different types: *financial*, *environmental*, and *cultural*.

At least one participant from all six groups mentioned and discussed a *financial* barrier to starting and conducting a multicenter study. Several participants stated that in general, they thought that obtaining funding was the biggest barrier to initiating and conducting a preclinical multicenter study. Many participants experienced this barrier, as only two study groups were fully funded. All participants from the unfunded or partially funded groups indicated that this was a barrier, one stating that lack of funding was a reason behind being unable to fully harmonize their protocols.

"We didn't have any funds for this… That's why it was challenging… it wasn't perhaps as harmonized as we'd liked" (7, D)

Participants of the fully funded groups stated that this was a barrier for them as well, as they felt that even the funding that they had was not enough to pursue their groups' future objectives. Interestingly, one participant indicated that funding was not a barrier as they managed to perform a successful multicenter study without external funding; they did, however, state that funding would be needed to expand their project to more than a "proof-of-concept" study for the use of this study design in preclinical research.

"*I think we need to go further. This was a pilot, a proof-of-concept, not the real thing.*" (2, A)

A second type of barrier that was discussed by multiple participants was *environmental*, which includes the environment across multicenter groups or within the participating laboratories' research space. The *environmental* barriers that participants mentioned included: differences in laboratory equipment and resources across sites, the geographical distances between sites (in reference to communication, transportation of materials, and being physically present at each site), and the general differences in lab environment such as temperature, air conditions, animal housing as well as institutional guidelines and regulations at each site (e.g. analgesia, method of animal sacrifice). Most of the discussion around *environmental* barriers centered on protocol development, and how these environmental aspects make it difficult to harmonize the protocols across sites.

"But a few things of course could not be easily harmonized… we had to adhere to legal regulations. For example, for analgesics which are also different across different European countries" (1, A)

"I mean everybody is working differently. For example, in our case everybody performs the stroke models differently, even if it's something super tiny… But yeah, we have to apply common standards now. I would say that these would be the hardest part" (3, A)

The third type of barrier discussed by participants from all six groups was classified as a cultural barrier. This barrier relates to both the culture of preclinical researchers and climate of the scientific community, and how this hinders the initiation, conduct and/or publication of multicenter studies. A *cultural* barrier faced by some participants changing their previously established experimental methods and techniques, in order to participate in the multicenter study. This resistance to change was a barrier to recruiting collaborators to join the multicenter group and/or to keep collaborators on track with using the agreed upon protocol. An additional *cultural*

barrier that was discussed by a few participants was the low publication yield and return on investment an individual researcher may receive from a multicenter study. The participants that discussed this saw it as a potential barrier to initiating future multicenter studies, as it may make it difficult to convince and recruit sites to join. However, they were not personally concerned about the limited opportunity to publish as a first or senior author.

"One barrier is of course, that the publication yield for the individual group is lower somehow. So you do an experiment, then you will be co-author. But if you do [a single-center] study, you will be the only author – the first or last author. So that's, I mean the scientific pay-off is lower somehow" (9, D)

Participants discussed the challenge of the scientific community not understanding the importance and value of preclinical multicenter studies. This referred to funding bodies, scientific journals, ethics committees, clinical trialists, and other preclinical researchers.

"It needs a change in attitudes from funding bodies. And I think also, the scientific community if I'm honest" (7, D)

Many participants stated that funding bodies did not have mechanisms in place to evaluate multicenter study grant applications and fund these types of studies, thus making it difficult to initiate and conduct one.

"It would change over-night if the funding bodies were willing to put money behind it and if they were much more willing to put emphasis on the need for replication rather than single studies" (7, D)

"If we don't have funding mechanisms for stuff like that it will remain an anomaly and only few of those trials will be conducted" (2, A)

Another *cultural* barrier that could impede the initiation of a multicenter study, and one which some participants faced in their respective studies, was that of research ethics board approval for multicenter animal experiments. This was the case for the studies that were conducted across Europe where different countries had different levels of specificity in ethics application, requirements, and regulations for animal experiments. One participant stated that they predicted that ethics approval may be one of the biggest challenges for future multicenter studies.

"A second unknown barrier, in fact this may be even a killer in some instances… is getting permissions for it. In different countries there are different ways of getting permissions for animal experiments… we have close to no experience how the regulatory authorities will react to study applications like this, because they may argue that we're repeating experiments. With a confirmatory [study], you show them that there already was such a study, and that gave some evidence to the efficacy of the compound… and now we really want to make sure. But to them this may appear excessive and unethical use of animals because we already have the evidence we need… So, if you ask me about hurdles or bottlenecks, this might be the one that kills it – at least in some countries or jurisdictions" (2, A)

Some participants felt that the broad community of translational scientists, trialists and preclinical researchers, misunderstood and did not fully appreciate the value of preclinical multicenter studies, and that this may hinder the popularization of this type of study design.

"So we need to clarify what the benefits of this are. And I think this concerns journals who publish stuff like that, I think this also concerns pharmaceutical companies who go on very expensive trials, which are based on very weak or not very robust evidence from the preclinical side" (2, A)

Enablers

All participants mentioned and discussed one or more types of enablers to the success of their own respective multicenter study or group, or to multicenter studies in general. There were four different types of enablers: *personal/social*, *skills and training*, *trust and transparency* and *financial*. Enablers were one of the most prominent discussion points within the interviews with participants, and many of the different types of enablers are highly linked to each other and with the barriers that were discussed previously.

All participants discussed *personal/social* enablers to their respective multicenter studies'. The majority of the discussion of personal and social enablers was centered on having had previously established relationships with other collaborators. All participants indicated that they had established relationships with the other sites' principal investigators.

"Well maybe it was successful because most of the people, or most of the PIs I was already working with and knew them. And that makes you confident. Because you know their work or the work that they are doing or did before" (3, A)

"I chose centers that had done animal studies before, with PIs that I knew well and would work in a collaborative fashion, and they had high quality and high functioning teams of people" (11, E)

Some relationships with collaborating principal investigators were more personal, as the collaborators were long-time friends, while other relationships were on a professional basis: as experts in the field they were known to each other. Other *personal/social* enablers that were discussed were aspects of the collaborators' personality. Some participants alluded to personality

traits that would be beneficial in facilitating a preclinical multicenter collaboration, such as being organized, a good multi-tasker and a team player.

"We're in academics already… as long as you're able to multitask as a person you're going to be okay" (5, C)

"So, this was a group of people who had a very similar mindset, and I think that that was one of the key elements of success" (2, A)

Another type of enabler that was mentioned by all participants was *skills and training*. *Skills* relate to the collaborating principle investigators' experience in their field and expertise with the animal models. Many participants indicated that being an established laboratory and a highly skilled investigator was essential to the success of their multicenter groups. Several stated that this not only enabled the study, it also made it more fundable and powerful. Some participants believed that an established laboratory was essential to multicenter studies, though others believed that this was not necessary but would be beneficial. *Training* included the preparation of lab personnel – technicians and students, prior to the start of the experiments. Many participants said that their study groups used a pilot trial, training sessions with technicians, or site visits where the personnel would visit each collaborating site to: see their procedures and equipment, practice the harmonized protocol, and resolve any discrepancies between sites. One participant stated that their group employed site visits for technicians, and that this was a critical feature to their collaboration's execution of the experiments and that future preclinical multicenter studies should have site visits.

Participants also identified *trust and transparency* as enablers. This included trusting in the skills, expertise and work habits of their collaborators; and being transparent with one another.

Multiple participants mentioned the aspect of *transparency* as an enabler. They stated that collaborators must be transparent with one another during the development and start of the study, throughout the course of the experiments, and during the dissemination of the results. Many participants indicated that they trusted their collaborators because they know them well personally or have worked with them before. Therefore, *trust* is linked with the enabler's *skills* and *personal/social* influences.

"We picked the three of us because we knew, and I knew that they would never decide to start short cutting… so I didn't have to worry about that. We started with absolute trust among the three of us. And there's a point at which you have to fundamentally trust your collaborators not to be intentionally misleading you, or not be sloppy in some way that you don't understand, so that the data their producing isn't valid" (4, B)

Site visits and training sessions involved some level of trust between collaborators, as the hosting sites need to trust their collaborators with intellectual property and must be able to share their methods between all technicians and participating sites.

The last type of enabler that was discussed by all participants was *financial* support – as seen through the discussion of funding as a barrier previously. Some participants indicated that not only obtaining funding, but a dedicated funding scheme within granting agencies for multicenter studies would be an important enabler.

"Well of course you need some money. So a dedicated funding scheme for multicenter trials for sure does help a lot, but it does not guarantee success" (1, A)

"The fact that there's money behind it really helps in terms of lubricating some of those collaborations" (6, C)

Objective 3: Extent of collaboration in preclinical multicenter studies

The third objective of this thesis was to evaluate the extent of collaboration in preclinical multicenter studies using Wood and Gray's theory of collaboration [27].

Elements of collaboration

Interview data showed that all multicenter groups demonstrated all six elements of collaboration discussed in the theory of collaboration: *convener, stakeholder autonomy, the interactive process, shared norms, rules, and structures, the action,* and *domain orientation*.

All participants identified a *convener* in their study. The *conveners* took on the role of initiating the study and identifying and bringing the collaborators together. The *conveners* that were identified in the studies appeared to be well connected individuals, well-respected in their field, and the most knowledgeable or interested in the preclinical multicenter study design within the collaboration. *Conveners* were the ones who "took charge" and got the "ball rolling", and their role was to convince and sell participating collaborators on the therapy being tested or on the multicenter design in general.

The second element identified was *stakeholder autonomy*. All participants indicated that joining and participating in their respective multicenter studies was on their own free will and that they retained decision making powers in some form. They all appeared to have been eager to join and participate in their multicenter studies. Overall, participants remained autonomous throughout the course of their study.

"And the ones that replied we followed up and discussed individually with each lab what they could do and couldn't do. That's really how it came about… these people did it out of generosity and collaboration" (7, D)

"I think that speaks to the collaborative nature of the project, that everybody has input, and everybody has a role to play, and everybody has a voice" (11, E)

The third element of the theory was *the interactive process*. All participants noted that the collaboration was an interactive process, in which the site principal investigators and other study collaborators interacted throughout the whole course of the collaboration. Some participants mentioned explicitly that the collaboration evolved over time, from the conception of the idea, to the development of the protocol, through the execution of the study, and finally to writing and disseminating the results.

"In the beginning we had multiple phone conferences, we exchanged several hundred emails, just to define actually how we want to perform this study. So of course, in the beginning the communication was very intense and was mainly in-between the PIs involved in that study. And this shifted afterwards a little bit when we came to the practical part of conducting the study, everything was defined" (1, A)

The fourth element – *shared norms, rules and structures*, discussed by all participants, appeared to vary in the level of formality across the six groups. Several participants from five of the six groups alluded to having shared professional *norms* among their multicenter group. The shared *norms* appeared to have emerged and been established prior to the start of the multicenter studies and consisted of skills and professional standards that are held by a high-quality preclinical scientist. These *norms* included using randomization and blinding techniques, power calculations for sample size, proper reporting of animal experiments, using appropriate controls, and not intentionally manipulating the data. The *norms* associated with high quality preclinical research were developed into and became explicit *rules* that were incorporated into the study design. Some participants expressed there was agreed upon formal *structures* within their respective

groups, specifically the processes in place for data sharing, quality control and protocol development, as well as dedicated meetings with other sites.

"We had monthly meetings every Tuesday, I think either the first or second Tuesday of the month. We had conference calls, and like I said, we would meet in person at a meeting… but if anything ever came up, [name] would initiate another conference call. We also had a manual of operations which got larger and larger and larger as we went through every drug… So, the manual of operations was very important" (5, C)

All participants identified *the action* as an element of collaboration – meaning that they identified a specific objective or goal for their respective multicenter collaborations. Though all participants stated that the goal was to test the efficacy of their respective therapies, there were some participants who expressed ancillary goals. Some participants discussed an additional goal of specifically performing a preclinical multicenter study as a proof-of-concept, while other additional goals were to mimic a clinical trial or to test the therapy in multiple models.

"So, the idea here was to put this on a broader basis, confirm it, increase external validity, put it through potentially other models, and well yeah make sure that the evidence from preclinical results is robust" (2, A)

The sixth element of the theory is *domain orientation*. All participants discussed the orientation of their own and their collaborators efforts and attentions to the action (goal/objective) of their multicenter study. The discussion of *domain orientation* was on action planning and decision making. Action planning involved protocol development, and decision making encompassed all aspects of study decisions from initiation to dissemination. All participants discussed both protocol development and study decisions. Study decisions included deciding which

collaborators/sites to invite, deciding to join the study, deciding on the model and level of protocol harmonization, logistic decisions, deciding how to interpret and write the results, and deciding which journal to submit to.

Overall, all six elements of the theory were identified by all of the preclinical multicenter study groups. According to the theory these six elements make up the definition of what collaboration is; therefore, this suggests that all six multicenter studies are in fact fully collaborative alliances.

Features of collaboration

Reasons for collaboration

Within the theory of collaboration, there are four central reasons for or precursors to collaboration between individuals and organizations: *high stakes and interdependencies*, *decreasing transaction costs*, *understanding the problem more fully*, and *maximizing efficiency of resource use* [27]. When applying this theory to preclinical multicenter studies, I defined the individual as the site principle investigator and the organization as their laboratory.

In the interviews, no participant alluded to or discussed *high stakes and interdependencies* between themselves/their lab and other principal investigators and their labs. Conversely, some participants indicated that preclinical investigators are highly independent. From this, it does not appear that *high stakes and interdependencies* are reasons for or precursors to collaboration in multicenter preclinical studies. Likewise, there were no participants that indicated that they engaged in the multicenter collaboration because they sought to decrease the transaction costs of performing a preclinical investigation, where transaction costs refer to any finance, resource, labor or time expenses. Most participants indicated that the multicenter collaboration increased the cost of conducting a preclinical study, and some participants stated that they were aware of

and planned for this before the initiation of their respective studies. From this, it does not appear that *decreasing transaction* costs is a reason behind preclinical multicenter collaboration.

Several participants discussed resources outside finances and funding such human resources (personnel, expertise) and material resources (animals, equipment), and how a multicenter study design can allow an investigator to make better use or gain access to these resources. For example, multiple participants discussed how a multicenter study increases the expertise and skill brought into the preclinical study that would not be there if it were a single-center study. There was also mention of how different labs engaging in the collaboration have different equipment available to them that are only accessible to others through collaboration. This may imply that the combination of expertise and equipment across otherwise independent labs could lead to *understanding the problem domain more fully* and in a transformative way through preclinical multicenter collaboration. Though this might not be the explicit reason to initiate a multicenter study, rather it could be a by-product or an ancillary advantage to performing one.

Additionally, several participants discussed the use of animals and how a multicenter study increases the sample size of their preclinical studies. Some participants discussed animal waste and waste of resources in single-center studies when they are underpowered. They believed that multicenter studies would make better use of animals and resources, as they are better designed and more robust than single-center studies. This suggests that *maximizing efficiency of resource use* may be a reason for engaging in a preclinical multicenter collaboration.

Outcomes of collaboration

Within the theory, there are four broad outcomes of which may be a consequence of; or they may increase or decrease because of collaboration. The two outcomes that can be a result of the

collaboration are: *introduction to bilateral and multilateral relationships*, *abandoning/reshaping and developing new skills*. The two outcomes that can either increase or decrease because of collaboration are: *environmental complexity and control*, and *transaction costs*.

Multiple participants indicated that the multicenter collaboration strengthened pre-existing relationships and introduced participants to new relationships with other. Therefore, it appears that *introduction to bilateral and/or multilateral relationships* are an outcome of multicenter collaboration. Likewise, most participants mentioned that they developed new skills and learned from the collaboration, some stating that they changed their techniques and aspects of their animal models. One participant stated that the quality of their preclinical work improved, and some indicated that their skills as a collaborator increased after being involved in their multicenter group. From this, it appears *reshaping old skills and developing new skills* is an outcome of collaboration in preclinical multicenter studies.

From the data, it appears that preclinical multicenter studies increase the *environmental complexity* and *reduce control*, as several participants indicated they had less control in the multicenter studies, the studies were more work, and there were greater logistical challenges than in single-center studies. Furthermore, because funding was a barrier and multicenter studies were more costly than single-center studies, it appears that increased *transaction costs* are an outcome of multicenter collaborations.

Chapter 4: Discussion

In this discussion chapter, I first address the three thesis research objectives. Following this, the remainder of the discussion is divided into three sections, where I discuss the results of the systematic review and the interview study separately, which is followed by a synthesized discussion of both studies.

Research Objectives

The first objective was to uncover the experiences and perceptions of individuals that have experience with preclinical multicenter studies. This objective was achieved through the combined findings of the systematic review and the interview study. The systematic review provided a preliminary evaluation of the aspects of interest through the narrative reports of the published preclinical multicenter studies. The interview study supported, solidified and expanded the initial fining of the systematic review through the interview with those who have participating in these studies. I found that researchers who have conducted and/or participated in a preclinical multicenter study held very positive opinions and views of these studies and had an overall positive experience with collaborating in these multicenter projects. From both the systematic review of studies and from the interviews, it appears that this type of study design is a useful tool and possibly an essential step in the translational process of novel biomedical therapies. Furthermore, it appears that these studies have a specific place and purpose in the translational pipeline.

The second objective was to identify any real and perceived barriers and enablers to preclinical multicenter studies and multicenter collaboration, from the perspectives of multicenter preclinical study stakeholders. Both the systematic review and the interview study address this

objective. In the systematic review, I extracted reports of potential barriers and enabler in the published preclinical multicenter studies; and I asked interview participants about the barriers and enablers they experienced and/or perceived for these types of studies. The most prominent aspect that either facilitated or hindered the initiation and conduct of preclinical multicenter studies was obtaining funding. In a similar vein, the culture and climate of the scientific community – particularly funding bodies, presented a significant barrier to future preclinical multicenter studies. Importantly, there were multiple enablers identified – both pragmatic enablers that can be reasonably implemented – such as training and site visits, as well as organizational enablers that cannot be directly realized – such as multicenter funding schemes.

The third objective was to evaluate the extent of collaboration in preclinical multicenter studies. The systematic review attempted to address this objective through the Degree of Collaboration assessment, but the limitations of this assessment did not allow for any compelling conclusions to be drawn. Thus, this objective was ultimately addressed through the interview study. It was found that the collaboration in preclinical multicenter groups represented in the interviews aligned with the definition of collaboration within Wood and Gray's theory of collaboration [27] and exhibited features that are present within the theory, suggesting that preclinical multicenter studies are true collaborative alliances and the extent of their alliance is a full collaboration.

4.1: Systematic Review

The first study of my thesis aimed to improve the awareness and knowledge of preclinical multicenter studies that have been conducted and published to date. I addressed this by performing a systematic review of preclinical multicenter studies in which I determined the differences and commonalities of these studies. Multiple narrative reviews and editorials have

been published on the topic of multicenter study design in preclinical research [14-17, 50]. In my thesis, I moved beyond these narrative pieces to synthesize characteristics and outcomes from all preclinical multicenter studies conducted to date. My results suggest that this is an emerging, novel and promising area of research, with thirteen interventional studies conducted to date, the majority published since 2015. Notable differences between the studies included clinical spheres, number of centers involved, and sample sizes. Similarities between studies included sources of funding (largely governmental), countries involved (three-quarters of centers were located in the USA), and the species of animal used (90% lab rodents).

A noteworthy finding is the discordance of results between previous single-center studies and subsequent multicenter preclinical studies. Five studies reported that their results confirmed previous single-center findings, six found no effect, and two found mixed effects. An explanation for why half of the multicenter studies reported null results, when the smaller primary studies conducted at single centers reported positive results is that this may reflect the increased sample size, methodological rigor, routine oversight and quality control of these multicenter studies. All of these factors would inherently lead to a more precise evaluation of the intervention's effects. A similar trend has been noted in clinical studies with decreased interventional effects, as studies move from single to multicenter studies [59]. This finding alludes to contentious issues of preclinical research: irreproducibility and poor generalizability of findings from single laboratories. As interventions are considered for translation, shifting away from traditional small-scale single-centered studies to a more methodologically rigorous, multicentric design may ultimately reduce translational failures and research waste.

Previous preclinical systematic reviews of single-center studies have found unclear or high risk of bias in most domains [60, 61]. In contrast, I found that multicenter studies generally had a lower

risk of bias in many more domains than typical single-centered studies. It has been established that practices such as blinding and randomization are known to affect the internal validity of both preclinical and clinical studies. Most studies in my review had low risk of bias in the three domains directly related to issues of randomization and blinding. Failing to use these methods represents a lack of rigor and is one of the suggested reasons behind failed translation [62].

Overall, the included studies had high completeness of reporting across most domains. Many of the items that all or most studies reported on were a part of the domains that address statistics and replicates. The items that studies reported at a lower frequency were specific to multicenter designs, such as indicating the number of participating centers in the abstract and identifying as a multicenter study in the title. The lack of multicenter or a synonym in the title is problematic when it comes to identifying preclinical multicenter studies in the literature. I faced this particular issue in my search for eligible studies, thus required multiple iterations of literature searches with updated search terms. One potential explanation for this finding is that guidelines and standards for multicenter studies are only just emerging, and there has yet to be any reporting recommendations specific to a preclinical multicenter design. Nonetheless, compared to previous reporting assessments of single-centered preclinical studies [63, 64], the completeness of reporting for the 13 preclinical multicenter studies was appreciably higher.

In general, I found a medium to high degree of collaboration. The domain of protocol development had the greatest proportion of studies assessed as high degree of collaboration, with no studies assessed as 'low.' Based on the degree of collaboration criteria a high score in the protocol development domain, this means that in most studies their protocol was developed *a priori* by all centers with or without a pilot study. Protocol execution had markedly lower scores, suggesting that centers may have found it difficult to maintain a high degree of collaboration

when the protocol was being executed. The degree of collaboration assessment has considerable limitations. The scoring was based solely on what was reported in the studies, and therefore is likely not a good measure of actual collaboration but rather an assessment of reporting practices. Additionally, there was only subtle distinction between the criteria used to judge whether each domain was given a low, medium or high DOC score – making it a highly subjective assessment. Importantly, the included studies were used to identify the population of interest for the interview study for Part Two of my thesis. Additionally, the systematic review also highlighted the aspects of preclinical multicenter studies that needed to be explored in greater depth in order to answer the underlying three objectives of my thesis. Specially, this information was used to aid in the development of the interview guide and the inductive analysis of Part Two of this thesis.

Strengths and limitations

A strength of this review is in the systematic synthesis of the published literature without limitation on the field or area of preclinical research, allowing for a more in-depth assessment of the state of this field of research. However, the scope of this review which allows for this in-depth, narrative report could also be seen as a limitation, as no quantitative results from the individual studies were assessed due to the heterogeneity of the studies. The application of rigorous inclusion criteria limited the eligible studies to interventional, controlled-comparison studies, which could omit valuable qualitative information that may have come from the excluded studies of non-controlled and/or observational designs. However, I purposely chose these criteria in order to focus on studies that might inform future interventional clinical trials. Another limitation of the review is that the degree of collaboration is solely based on what is reported in the studies, thus the insights gained into the nature and degree of collaboration within

the included studies was limited. It is with this limitation in mind that the qualitative interview study was performed as the second half of this thesis.

4.2: Interview Study

The interview study consisted of in-depth, semi-structured interviews with 12 researchers who participated in and/or conducted the preclinical multicenter studies. In these interviews, I investigated the experiences, views and opinions of preclinical multicenter studies, uncovered barriers and enablers to initiating and conducting such studies, as well as explored collaboration between stakeholders participating in such multicenter projects.

Similar to what was found in the results of the systematic review, the structures and processes of the preclinical multicenter groups that were represented in the interviews varied across groups. Although there were distinct logistical differences between the groups, participants felt that their respective multicenter studies were in general successful and were worthwhile to their career and to their fields of study. Since all the multicenter studies were considered "successes" by participants regardless of their differences, it can be inferred that there is no prescriptive structure or "one-size fits all" approach to multicenter studies in regard to the basic logistical features. However, there were certain organizational and structural features that were shared across multicenter groups, and/or were suggested as potential enablers for future preclinical multicenter studies. This included a coordination center, a data analysis center(s), and an external review committee.

A prominent theme that emerged from the interviews was that of appropriate methodology and design of preclinical multicenter studies. A similar feature between the six groups and a proposed enabler was that preclinical multicenter studies should mimic a multicenter *clinical*

trial. This suggests that preclinical multicenter trials could use the well-designed clinical trial as a template in the planning of future studies. Specifically, preclinical studies should adopt features that are well established benchmarks in clinical trials such as proper blinding and randomization practices, external (non-experimental site) regulation and oversight, agreed-upon operating procedures and manual of operations, data sharing, and quality control. Of these practices and features, the one discussed in the greatest depth was the harmonization of protocols across sites. There were three different approaches to protocol harmonization across the six groups: fully harmonized, multiple models, and not harmonized. The one group that did not have a harmonized protocol faced the greatest challenges and shared that the lack of harmonization was the only regret they had about their study. The groups that had fully harmonized or multiple model protocols seemed to have very few challenges and believed that their level of protocol harmonization was suitable for the goals of their respective studies. This suggests that preclinical multicenter studies should not employ a non-harmonized protocol, and instead should share either a completely harmonized protocol across participating sites or have harmonized procedures with multiple models. Furthermore, the process of protocol development should be a shared task between all sites, as this is what occurred in the five groups with fully harmonized or multiple models and was a suggested enabler to preclinical multicenter studies.

The choice between these two harmonization approaches may be best determined jointly by all participating sites and is dependent on what the multicenter study is trying to achieve and the nature of the disease model of interest. Several of the participants whose groups used a fully harmonized protocol discussed the “replication crisis” in their field of study, in that there had been many potential therapies tested in previous preclinical studies using the exact same animal model, but the efficacy of the therapies was not consistent across independent preclinical studies.

Some participants from groups that used multiple model harmonized protocols indicated that there were few promising therapies for the disease of interest - potentially because the preclinical models do not replicate the clinical condition and/or the condition is complex and multifaceted. This could indicate that a fully harmonized protocol should be used when the disease domain of interest has a well-established animal model as well as promising therapies that cannot be replicated and the findings cannot be reproduced across labs. A possible circumstance for when a multiple model protocol could be used across participating sites is when there is little to no consensus on the best animal model of the disease, and when the clinical condition is highly complex and/or difficult to replicate in a preclinical setting. This situation would speak to the generalizability of the previous research in that field, rather than the reproducibility of a potential therapy.

The choice in the type of protocol may also depend on the previous work that has been done in the field or on the therapy of interest, in that the design of the multicenter study's purpose is to remediate the issues that have hindered the therapy's previous attempts of translation. If previous work has been underpowered, then the most critical feature of the multicenter study is the increase in sample size with multiple sites. This increase in sample size would only be realized if all sites perform the exact same protocol – essentially the same study. Alternatively, if the proposed major issue of translation is that previous research was unable to mimic the clinical setting in an animal model, it would be beneficial to use a triangulation method in which the same therapy is assessed using a combination of methods or models.

The type of protocol harmonization undertaken may not always be a choice and is perhaps influenced by logistical matters such as available funds and expertise of the participating laboratories. Having multiple models that are adequately powered would likely require a larger

number of animals than if only one model was used; therefore the multicenter studies may not have the funds to apply multiple models. A multiple model protocol may also not be possible if the participating sites all perform the same model and do not have the resources or expertise to incorporate another one.

Many of the participants strongly believed that the multicenter design should be used explicitly for confirmatory preclinical research or "validation trials". This explains why some participants had difficulty in directly comparing single centered to multicentered preclinical research, as they felt that they were applied to different research contexts. It appeared that participants felt that single centered research should be used for exploratory and discovery research but was not sufficient in determining whether a novel therapy should be tested in human volunteers. Furthermore, participants felt that a multicenter study would be wasteful or unnecessary in exploring biological mechanisms or performing a drug discovery study. Participants stated that only once a therapy is established and there is reason to believe that it could be translated to a first-in-human trial, should it be evaluated in a preclinical multicenter study. A notable finding is that the preclinical multicenter study design has a distinct place within the pipeline of translational research and a definitive purpose and reasoning for why it is used. In general, it appears preclinical multicenter studies should be conducted when a therapy has been evaluated in previous single centered research and is being considered for a first-in-human trial. The general purpose of a preclinical multicenter study is to confirm and validate a novel therapy through a rigorous and controlled design which has the potential to evaluate replicability and generalizability.

From the interviews I found that participants had very positive experiences with, and opinions of, preclinical multicenter studies. Participants generally held these positive attitudes due to the

benefits they experienced from the collaborative effort – mainly learning, building relationships, and advancing their career. Learning of new knowledge, skills and techniques was discussed extensively and it appears to be an ancillary benefit that is achieved from a collaborative venture such as a preclinical multicenter study. A preclinical researcher isolated in their own laboratory would not be exposed to other researchers' methods and expertise to the same extent as in a multicenter study. This level of sharing, and the resulting knowledge gained, is a benefit unique to multicenter projects as opposed to single-centered ones. Similarly, researchers collaborating on a multicenter study can be introduced to and develop relationships with other experts that they may not otherwise have been if they worked solely on single-centered projects. A third ancillary benefit, career advancement, is likely due to the combination of the two benefits of learning and relationship building, as well as the publications and the publicity that comes from them. Participants stated that collaborating with their multicenter groups led to important publications, recognition, and thus advanced them in their career. This could be because the preclinical multicenter research design aims to address alarming issues such as translational failure and lack of reproducibility in many areas and fields of biomedical research, which have received attention from top-tier journals and funders [65, 66]. Additionally, the preclinical multicenter design is a novel approach that alters the traditional archetype of preclinical research. Both matters may bring about discussion and debate by preclinical, clinical, and translational scientists, which could lead to attention and exposure for those collaborating on a multicenter study.

Additional notable findings were the several barriers to preclinical multicenter studies that were uncovered. The most palpable barrier for preclinical multicenter studies is funding for these projects. It is natural that this should be a barrier as it is a challenge experienced in scientific research projects in general and is not a barrier that is completely unique to preclinical

multicenter studies. However, this barrier may be more significant in multicenter studies as they are more expensive than single-centered studies, and there are few dedicated funding mechanisms in place for these types of projects. This lack of funding mechanism leads into a second notable barrier that was uncovered in the interviews: the culture and climate of the scientific community. Participants felt that the scientific community did not understand the importance and reasoning behind preclinical multicenter studies, and that this may hinder the initiation and conduct of future multicenter studies. If funding bodies do not appreciate these types of studies, they will not be funded. Likewise, if research and ethics boards do not understand what these studies are trying to achieve, they may not approve the animal experiments. Moreover, if scientific journals do not recognize the novelty and contribution of preclinical multicenter studies to the field of science, they may not be as widely disseminated.

On a smaller scale, the culture and norms of the individual preclinical researchers may pose a barrier in multiple ways. First, low publication yield may be a barrier for recruiting collaborators to join a multicenter group. This could be due to the pressure of "publish-or-perish" an individual researcher may experience, which is rooted in academic science [67, 68]. Furthermore, when a preclinical principal investigator collaborates on a multicenter study instead of single-center study, they are often foregoing the right to be the first author or senior author on the work. This is a barrier as scientists are rewarded for first and senior author publications more than for co-authored publications. Additionally, preclinical scientists traditionally work in isolation – having full control of how they conduct their research. Participants stated that they occasionally found it challenging to harmonize protocols because the collaborators were all accustomed to their own methods. If researchers are inflexible and uncompromising, they are unlikely to change their

methods in order to harmonize a protocol across multiple sites and would be less willing to join a multicenter collaboration.

Several enablers to preclinical multicenter studies were also identified through the interviews. After considering the identified enablers I synthesized them into two categories: pragmatic enablers and organizational enablers. The pragmatic enablers are those the participants themselves employed in their multicenter groups and/or believed that they should be implemented in future preclinical multicenter studies. As the name suggests, pragmatic enablers are realistic and appear to be achievable – they can also be implemented directly in the multicenter study by participating collaborators. These pragmatic enablers include: experimenter training and site visits, regular and structured communication channels with all participating sites, employing a site point-person to manage the logistics of the multicenter group, being transparent and open with all collaborators, and criteria for site selection – where recruited sites should be based upon 1) the skill and expertise of the laboratory and 2) previously established relationship and trust in the site PI.

The enablers that were categorized as organizational enablers are less practical, and potentially unrealistic and infeasible in that collaborators do not have direct control in ensuring these enablers are in place prior to conducting their own preclinical multicenter study. These large-scale enablers include a funding scheme specifically for preclinical multicenter studies, mechanisms for multi-site research and ethics board applications for animal experiments, a change in scientific culture where preclinical research is more conducive to collaborations and that there is better knowledge of and appreciation for the preclinical multicenter study design. The large-scale enablers that were discussed were largely in the context of addressing the main barriers participants themselves experienced and anticipated for the future.

It is also interesting to note that many of the participants appeared to have limited to no knowledge of other preclinical multicenter groups or of studies that had been conducted before or after their own respective studies had been published. Several participants stated that their preclinical multicenter study was the first of its kind, when in fact there were such studies published before their own. This could be due to the fact that this is such a novel approach to preclinical research and practices for naming and reporting these types of studies have not been fully defined in the literature. It is understandable that preclinical researchers are not fully aware of other multicenter studies that have been published, as I myself had difficulties in identifying these studies with a systematic search in Part One of my thesis. The participants' lack of awareness combined with the challenges in identifying these studies provides insight into the deficiencies in indexing and characterizing preclinical multicenter studies, as well as participants' belief that there is a need for a better appreciation and understanding of these studies by the translational research community.

As per the third objective of my thesis, I used six elements of Wood and Gray's theory to evaluate the extent of collaboration within preclinical multicenter projects. The six elements were identified in all multicenter groups, suggesting that preclinical multicenter groups conform to the definition of collaboration and are thus true collaborative alliances.

It was found that preclinical researchers do not engage in a collaborative alliance due to high stakes and interdependencies. High stakes suggest that the situation involves significant risk and the processes and decisions are of great importance, while interdependencies suggests reliance's that are created because organizations (laboratories) possess and control necessary resources. Laboratories are not always in complete control of their own resources as they rely on grants and awards to fund their research and must adhere to the policies of the larger organizations that they

are a part of (academic hospitals, pharmaceutical companies, and universities). Additionally, preclinical laboratories often do not depend on, or have direct influence over other laboratories. Reducing transaction costs was also not a reason behind preclinical multicenter collaboration, as transaction cost reduction suggests that the expenditure would be less in a multicenter collaboration than they would in a single-centered research project.

It is unreasonable to assume that because these are not precursors to multicenter collaboration, that these are also not conditions researchers regularly confront or are challenged with in the preclinical domain. Rather, high stakes, interdependencies and high transaction costs could be challenges for preclinical researcher, but they are not the problems that the collaborators are trying to address. If this is the case, it could be that multicenter collaboration increases that stakes and interdependencies; and the results showed that transaction costs did increase due to collaboration.

The theory does not state that these conditions must be present for a collaboration to be initiated, but instead are potential precursors. Therefore, though high stakes and interdependencies and reducing transaction costs are not reasons for preclinical multicenter collaboration, this does not mean that the theory of collaboration does not support these types of alliances.

Both understanding the problem more fully and maximizing efficiency of resource use appear to be reasons for collaborating in preclinical multicenter studies. Understanding the problem more fully may be the most central reasoning for multicenter collaboration and that is the primary reason why the participating researchers are performing the experiments and collaborating. The argument that multicenter studies maximize efficiency of resources is that multiple underpowered, and sometimes flawed single-centered studies are not sufficient for confirmatory research, therefore, animal lives are unnecessarily wasted, laboratories exhaust their own time

and resources, and no new insights are learned about the potential efficacy of the new therapy. Multicenter studies may avoid these failings and reduce research waste of single-centered studies. Additionally, efficiency refers to how well something is done. From the interviews it appears that multicenter studies are more robust and rigorous than single-centered studies, therefore, though individual collaborating labs expend more resources, overall these resources are put to better use, thus the efficiency of these resource are being maximized.

Contributions to the theory of collaboration

My research supports Wood and Gray's theory of collaboration in the setting of preclinical multicenter studies. This suggests that this theory is appropriate for analyzing collaboration in this area of study. Therefore, if the construct of collaboration is to be studied further in preclinical research, I suggest applying Thomson and Perry's developed framework of five dimensions, to determine whether the nature of collaboration is comparable between the field of biomedical research and public administration. Furthermore, Thompson, Perry and Miller [69] go on to conceptualize and test to construct validity of collaboration, creating a mathematical model to measure the five dimensions of collaboration, and thus a collective quantitative measure of overall collaboration. By applying this model, this would allow for a more precise measurement of the extent this collaboration in preclinical multicenter alliances, as the validated mathematical model would produce quantitative scores of collaboration rather than an evaluation based on judgement. This would also contribute in the effort to further cross-validate this model. Cross validating the model using data from the preclinical setting would increase confidence in the model's accuracy the biomedical field and would increase generalizability to other fields of research seeking to measure collaboration.

The context and setting examined in my research differed from those of the original case-studies that formed the theory, as well as those that were used to develop the theory into a framework and then a mathematical model. Though my search of literature on collaboration was not systematic, it appears that my thesis is the only instance of the theory being applied to the preclinical research domain. Thus, my research is an important and novel contribution to the theory of collaboration.

Barriers, enablers and incentives to collaboration in clinical trials

Overall, I did not find any empirical research into the nature of collaboration between stakeholders in the preclinical biomedical domain, thus I cannot make an evaluation as to how my results compare to other existing research. Several qualitative studies within the clinical domain have investigated aspects such as structural features and processes, barriers, enablers and best practices for clinical research and implementation research [70, 71]. Much of the work focussed on identifying the barriers and enablers to clinical trials, concerns the recruitment and retention of patients for the trials [72, 73], while implementation research focuses on studying the best methods and processes for implementing novel guidelines and interventions to practice in the clinical setting [70, 74]. Comparing these processes, barriers, and enablers to those that I identified in my thesis provides little value, as patient recruitment for the research subject and process within the clinical domain are not applicable to the preclinical stage of translational research. There is another body of research that focuses on physician participation and engagement in clinical trial research [75], which may provide a more valuable comparison, as preclinical investigators need to be recruited to participate in preclinical multicenter projects.

Barriers to physician and clinical researcher recruitment and engagement that have been identified in previous research include increased workload and time demands for research, the

burden of educating patients, adhering to patient follow-up schedules, and the lack of personnel and research support staff [74-76]. The barriers of increased time and work demands, patient education, and follow-ups for participating physicians are not applicable to preclinical principle investigators as the primary role of a preclinical scientist is to conduct research, and patients are not involved in the preclinical phase. Lack of personnel and support staff, which is essentially a lack of funding, is pertinent to both preclinical as well as clinical research, as this was identified as a barrier in preclinical multicenter studies. Enablers and incentives for clinical research participation include building personal connections between research leads and the recruited physicians, selling the research objectives and methods to physicians, the opportunity for professional development, contributing to improved practice and knowledge, regular meetings during trial, and training with protocol [75-77]. The enablers and incentives of clinical research participation appear to be more pertinent to preclinical multicenter studies, as many of these enablers and incentives are consistent between the two areas of research.

Collaboration is a valuable strategy that organizations can use to cope with the complexities and challenges within their environment[27]. As such, many organizations and stakeholders in various settings use collaboration to address complex problems they have not been able solve independently[33]. Like any tool or strategy, it is important that collaboration is understood in order to derive the most value from it and use it effectively. Specifically, effective collaboration may improve team function, build good relationships, reduce resource use, and boost learning; all of which lead to an increased possibility of solving the problem that brought collaborators together [33]. The concept of collaboration within biomedical sciences does not appear to have been explored to the same extent as in other fields such as business and public administration. My exploration of collaboration in preclinical multicenter studies adds new insight into the

nature of collaboration within the domain of biomedical research and is therefore an important step in improving the collaborative process in this field.

Strengths and limitations

A strength of this interview study is that it applied both in-depth inductive and theory-informed deductive methods. An additional strength is that members from almost all the groups that have conducted a preclinical multicenter study were interviews and were accounted for in this study (one group couldn't be contacted; another group's members were all retired or deceased). Another strength was that it was informed and supported by the systematic review of Part One, where it aided in the development of the interview guide and analysis, thus making the interview process more efficient. However, this interview study has several limitations that should be considered.

The first limitation of the study is that the sample size was limited to 12 participants, from six independent preclinical multicenter groups. Only eight preclinical groups had been identified, but participants from two of them could not be contacted for an interview. This could limit valuable insight that may have been procured from these two groups. Additionally, the findings of this study may not be generalizable to all researchers conducting a preclinical multicenter study as participants' demographics were restricted to five fields of research and four countries. These limitations are unavoidable as the population of interest is limited and a 100% response rate was not expected.

Another potential limitation is that all interviews were conducted in English, and English was not the first language for several of the participants. Though I took care to use plain language, speak slowly and repeat questions when necessary, it is possible that some participants may have misunderstood questions or may not have been able to respond to their full capacity had the

interview been conducted in their native language. Another potential limitation is that the interview participants have conducted preclinical multicenter studies, which may bias their opinions, views and attitudes. Participants could potentially be acting as advocates for this study design and hold overly positive views towards clinical multicenter studies. In future research, it may be valuable to interview a sample of preclinical researchers that have not conducted or participated in a preclinical multicenter study but have knowledge of such studies.

4.3: Synthesis of systematic review and interview study findings

In this section, I combine and compare the results from the systematic review and the interview study in a synthesized discussion; and make recommendations for the conduct and reporting of future preclinical multicenter studies based on my results. I discuss the main findings of the systematic review, and then explain how the results from the interview support these findings. First, I discuss the observed low risk of bias and high completeness of reporting, and how the interview study results explain this finding. Following this I discuss the null results of the multicenter studies and the inconsistency from the previous single-center positive results, and how this was something interview participants did not expect. Lastly, I emphasize how the barriers and enablers extracted from the systematic review are supported by those that were uncovered in the interview study.

From the systematic review, it was found that the preclinical multicenter studies generally had low risk of bias in many of the domains – many of which address preclinical practices of blinding and randomization. Additionally, the included studies had high completeness of reporting across most domains, with high completeness of reporting on items focussed on statistics and replicates. After considering what was uncovered and discussed in the interviews, these results are not surprising. One of the findings from the interview study was that preclinical

multicenter studies attempt to mimic a multicenter clinical trial, as well increase the robustness and rigor in the preclinical phase. Participants explicitly mention performing proper blinding and randomization practices in their respective studies and state that this is an underlying goal of preclinical multicenter studies in general.

Another finding of the systematic review was that the results of the previous single-center studies did not align with the results of the preclinical multicenter studies to which they supported. All single-center studies found positive results in the tested therapies, while only five of the thirteen therapies were positive in the multicenter studies. When study outcomes were discussed by interview participants, many expressed that they did not expect their study to produce null results for the therapy they were testing, some stating that their surprise was due to the previous positive results of single-centered studies. Though participants voiced surprise with this inconsistency in their results, it is something that could have been anticipated. In the clinical domain, it is typical for multicenter trials to demonstrate no effect from the treatment, when previous findings in single-center trials found the treatment efficacious [59, 78]. Scientists who have observed and studied this issue in clinical trials, suggest this likely reflects the increased rigor of multicenter studies, and that the results from single-center trials should be used with caution in decision making [79]. While further research and empirical evidence is needed to determine whether this phenomenon occurs in preclinical research, my findings suggest that multicenter studies show a smaller treatment effect than single-center studies.

The barriers that were identified in the systematic review are consistent and are supported by the barriers identified in the interview study. Protocol development and issues in the early stages of the multicenter projects were challenges that were identified in the systematic review. Similarly, harmonizing the protocol was a major theme that emerged in the interviews and was discussed

in-depth by many participants. Both of these findings suggest that protocol harmonization is a critical element of consideration when designing a preclinical multicenter study.

Recommendations for preclinical multicenter studies

After performing an assessment of the completeness of reporting of the published preclinical multicenter studies, uncovering barriers and enablers to such studies, and exploring the experiences of those who have participated in a preclinical multicenter group, I provide recommendations for future preclinical multicenter studies in Box 1. It should be acknowledged that these recommendations are of my opinion and are based solely on results of just 13 preclinical multicenter studies that have been published, and interviews with 12 preclinical researchers who have conducted these studies. Therefore, the recommendations may not be generalizable to all future preclinical multicenter studies.

Box 1 –Recommendations for preclinical multicenter studies

Conduct	**Reporting**
Planning stage - Conduct rigorous budget planning and funding allocation across centers - Consider industry involvement as costs will be substantial - Collaboratively develop protocol with exact details across all participating centers - Evaluate if multiple animal models of the disease are needed to address translational issues of previous research - Rigorously standardize experimental conditions across all participating centers - Involve animal ethics boards and other regulatory stakeholders early in study protocol development - Conduct a pilot study involving all participating centers - Conduct laboratory technician training sessions during and after the pilot studies in order to assure lab techniques will be performed correctly - Perform site visits with participating centers to visualize exactly what is being done at each site *Experimental stage* - Conduct regular progress meetings throughout the study's entirety with all participating centers - Assign a center to be responsible for study coordination and data analysis	*Basic elements* - Indicate that the study is a preclinical multicenter study in the title and abstract - Provide a description of how the study is a multicenter study - Use consistent terminology (i.e. preclinical multicenter study) *Design* - Explicitly state the number of participating centers (in the abstract and/or methods) - Indicate the exact role of each participating center - Indicate what aspects of the study design were shared between centers - Indicate what aspects of the study design differed between centers - At each center, indicate the number of animals used, outcomes assessed, and analyses performed - Explicitly state the number of animals in each experimental group and the exclusions and losses of animals at each center with explanations - Report basic elements according to the ARRIVE guidelines or the NIH preclinical reporting guidelines [13, 80]

Chapter 5: General Conclusions

This thesis employed rigorous methodology to evaluate the emerging field of preclinical multicenter studies. I found that preclinical multicenter studies aim to address the lack of preclinical study result replication and possible deficiencies in the preclinical stage of biomedical research. Even though a small number of these studies have been conducted thus far, I identified several barriers and enablers and important features of these studies from reports within the literature, and through the experience of those who have conducted them. Specifically, the important considerations for future preclinical multicenter studies include harmonization of protocols, funding for studies, the decision to employ one or multiple animal models, the inclusion of non-experimental centers, site visits and training sessions, the selection of participating sites and principal investigators, and the collaborative nature of these studies.

Despite there only being 13 multicenter studies published to date, the systematic review demonstrates the rigor of a multicentric study design, as it was found that these studies employed extensive quality control and study oversight. This potentially led to the observed lower risk of bias in many key domains and higher completeness of reporting compared to typical single laboratory studies. The review provides evidence that preclinical multicenter studies may be used as a tool to inform further development and future studies of promising clinical therapies at the confirmatory, preclinical stage of biomedical research.

The interview study further supported that preclinical multicenter studies may be used as an effective approach in the translation of novel therapies, as many experts in the field of preclinical research believe that these studies are a critical part of the translational pipeline. Specifically, multiple participants stated that preclinical multicenter studies should be used for confirmatory studies that seek to validate the efficacy of a therapeutic intervention before it is translated into a

first-in-human trial. Furthermore, it was found that these types of studies involved true collaborative alliance between multiple stakeholders. Knowing that these are collaborative projects is important in the planning of future studies, as stakeholders interested in performing a multicenter study may borrow from the extensive literature and research that aims at improving the collaborative processes and strengthening collaborative alliances [81, 82].

Future research

An important next step to improving the design and increasing the adoption of preclinical multicenter studies would be to identify consensus-based practices for these studies. To do this, I suggest performing a Delphi exercise involving experts in the field biomedical sciences, translational research, and preclinical scientists that have conducted a multicenter study. The best practices would include guidelines for conducting and reporting preclinical multicenter. Before these guidelines can be developed, additional research should be conducted to address important aspects of preclinical multicenter studies that I identified in my research but require greater clarity. Questions that should be addressed include: what type of protocol is best – a fully harmonized method or a triangulation method? And; for what situation or research goal would each protocol be most optimal? Furthermore, what number of experimental centers is optimal, or what number is best for particular settings or conditions? Addressing these questions would require empirical evidence, using the different preclinical multicenter studies to conduct case-studies and through computer simulation research[83].

In the systematic review of preclinical multicenter studies, I made an assessment of the completeness of reporting of such studies using a modified checklist from various sources. This checklist was sufficient for my thesis despite notable limitations – primarily in reporting items that were specific to preclinical multicenter studies, such as indicating “multicenter” (or

synonym in the title), or the number and roles of participating centers. If multicenter studies are to be used more extensively in preclinical research, it would be valuable to have a reporting checklist for these types of studies. As such, I propose that a consensus-based reporting checklist be constructed through appropriate methodology such as a Delphi exercise involving experts in the field of quality of reporting, as well as those mentioned above. Delphi exercises have been used in the past to develop reporting guidelines in different areas of expertise such as health research and systematic reviews. The EQUATOR (Enhancing the QUAlity and Transparency Of health Research) Network suggests that such a consensus process is crucial in developing a reporting guideline [84].

Likely the most important concern to address to with future research is: how do preclinical multicenter studies actually affect translation? The premise of conducting preclinical multicenter studies is that they more robust than single-center studies and are better suited to confirming if novel therapies should enter clinical trials. Whether or not this is true should be the central concern of future research as it appears these studies are more time consuming and costly than single center-studies. Therefore, it would be wasteful and less cost-effective if the multicenter design is not superior to the single-center design for confirmation studies. Again, empirical evidence is needed to first evaluate the quality and strength in evidence of multicenter as compared to single-center studies; as well as evidence on how multicenter studies affect the translation of novel therapies further into the translational pipeline.

Conclusion

Advancement in the ability to treat diseases and medical conditions is essential in the improvement of human health and wellbeing. These advancements will not be realized if the novel biomedical discoveries of basic research cannot be translated into viable clinical therapies.

Therefore, there is an urgent need to improve the process of advancing therapies from the preclinical to clinical stage of medical research. The work of my thesis aimed to address this concern through the exploration of preclinical multicenter studies – a possible solution to improving preclinical to clinical translation. With the better understanding of how preclinical multicenter studies have been and can be applied along with the identified the barriers and enabler to these studies, this will allow for more effective planning and execution of future multicenter projects. This will support greater efficiency in preclinical multicenter studies, which could increase the chance of success in clinical trials, and thus lead to safe and effective medical treatments.

My research suggests that preclinical multicenter studies may be a potential way in which we can improve the process of translation, thus reducing research waste and delivering needed therapies to people burdened with disease worldwide.

References

1. Grimshaw JM, Eccles MP, Lavis JN, et al. Knowledge translation of research findings. *Implement Sci* 2012; 7: 50. 2012/06/02. DOI: 10.1186/1748-5908-7-50.
2. Collins FS. Reengineering translational science: The time is right. *Sci Transl Med* 2011; 3. DOI: 10.1126/scitransmed.3002747.
3. DiMasi JA, Grabowski HG and Hansen RW. Innovation in the pharmaceutical industry: New estimates of R&D costs. *J Health Econ* 2016; 47: 20-33. 2016/03/02. DOI: 10.1016/j.jhealeco.2016.01.012.
4. Hanney SR, Castle-Clarke S, Grant J, et al. How long does biomedical research take? Studying the time taken between biomedical and health research and its translation into products, policy, and practice. *Health Res Policy Syst* 2015; 13: 1. 2015/01/02. DOI: 10.1186/1478-4505-13-1.
5. Settleman J and Cohen RL. Communication in Drug Development: "Translating" Scientific Discovery. *Cell* 2016; 164: 1101-1104. 2016/03/12. DOI: 10.1016/j.cell.2016.02.050.
6. Alberts B, Cicerone RJ, Fienberg SE, et al. SCIENTIFIC INTEGRITY. Self-correction in science at work. *Science* 2015; 348: 1420-1422. 2015/06/27. DOI: 10.1126/science.aab3847.
7. Perry CJ and Lawrence AJ. Hurdles in Basic Science Translation. *Front Pharmacol* 2017; 8: 478. 2017/08/05. DOI: 10.3389/fphar.2017.00478.
8. Smith C, Baveja R, Grieb T, et al. Toward a science of translational science. *J Clin Transl Sci* 2017; 1: 253-255. 2018/04/17. DOI: 10.1017/cts.2017.14.
9. Drolet BC and Lorenzi NM. Translational research: understanding the continuum from bench to bedside. *Transl Res* 2011; 157: 1-5. 2010/12/15. DOI: 10.1016/j.trsl.2010.10.002.
10. Morris ZS, Wooding S and Grant J. The answer is 17 years, what is the question: understanding time lags in translational research. *J R Soc Med* 2011; 104: 510-520. 2011/12/20. DOI: 10.1258/jrsm.2011.110180.

11. Contopoulos-Ioannidis DG, Ntzani EE and Ioannidis JPA. Translation of highly promising basic science research into clinical applications. *The American Journal of Medicine* 2003; 114: 477-484. DOI: 10.1016/s0002-9343(03)00013-5.

12. Bath PM, MacLeod M and Green AR. Emulating multicenter clinical stroke trials: a new paradigm for studying novel interventions in experimental models of stroke. *Int J Stroke* 2009; 4: 471-479.

13. Landis SC, Amara SG, Asadullah K, et al. A call for transparent reporting to optimize the predictive value of preclinical research. *Nature* 2012; 490: 187-191. 2012/10/13. DOI: 10.1038/nature11556.

14. Llovera G and Liesz A. The next step in translational research: lessons learned from the first preclinical randomized controlled trial. *J Neurochem* 2016; 139 Suppl 2: 271-279. 2016/11/01. DOI: 10.1111/jnc.13516.

15. Dirnagl U and Fisher M. REPRINT: International, multicenter randomized preclinical trials in translational stroke research: it is time to act. *Stroke* 2012; 43: 1453-1454. 2012/04/27. DOI: 10.1161/STROKEAHA.112.653709.

16. Fernandez-Jimenez R and Ibanez B. CAESAR: one step beyond in the construction of a translational bridge for cardioprotection. *Circ Res* 2015; 116: 554-556. 2015/02/14. DOI: 10.1161/CIRCRESAHA.115.305841.

17. O'Brien TJ, Ben-Menachem E, Bertram EH, 3rd, et al. Proposal for a "phase II" multicenter trial model for preclinical new antiepilepsy therapy development. *Epilepsia* 2013; 54 Suppl 4: 70-74. 2013/08/09. DOI: 10.1111/epi.12300.

18. Anello C, O'Neill RT and Dubey S. Multicentre trials: a US regulatory perspective. *Stat Methods Med Res* 2005; 14: 303-318. 2005/06/23. DOI: 10.1191/0962280205sm398oa.

19. van der Worp HB and Sandercock PA. Improving the process of translational research. *BMJ* 2012; 345: e7837. 2012/11/28. DOI: 10.1136/bmj.e7837.

20. Jones SP, Tang XL, Guo Y, et al. The NHLBI-sponsored Consortium for preclinicAl assESsment of cARdioprotective therapies (CAESAR): a new paradigm for rigorous, accurate, and reproducible

evaluation of putative infarct-sparing interventions in mice, rabbits, and pigs. *Circ Res* 2015; 116: 572-586. 2014/12/17. DOI: 10.1161/CIRCRESAHA.116.305462.

21. Boltze J, Wagner DC, Henninger N, et al. Phase III Preclinical Trials in Translational Stroke Research: Community Response on Framework and Guidelines. *Transl Stroke Res* 2016; 7: 241-247. 2016/06/15. DOI: 10.1007/s12975-016-0474-6.
22. Llovera G, Hofmann K, Roth S, et al. Results of a preclinical randomized controlled multicenter trial (pRCT): Anti-CD49d treatment for acute brain ischemia. *Sci Transl Med* 2015; 7: 299ra121. 2015/08/08. DOI: 10.1126/scitranslmed.aaa9853.
23. Maysami S, Wong R, Pradillo JM, et al. A cross-laboratory preclinical study on the effectiveness of interleukin-1 receptor antagonist in stroke. *J Cereb Blood Flow Metab* 2016; 36: 596-605. 2015/12/15. DOI: 10.1177/0271678X15606714.
24. Wahlsten D, Metten P, Phillips TJ, et al. Different data from different labs: lessons from studies of gene-environment interaction. *J Neurobiol* 2003; 54: 283-311. 2002/12/18. DOI: 10.1002/neu.10173.
25. Reimer KA, Jennings RB, Cobb FR, et al. Animal models for protecting ischemic myocardium: Results of the NHLBI Cooperative Study. Comparisons of unconscious and conscious dog models. *Circ Res* 1985; 56: 651-665. DOI: 10.1161/01.RES.56.5.651.
26. Chung KC, Song JW and Group WS. A guide to organizing a multicenter clinical trial. *Plast Reconstr Surg* 2010; 126: 515-523. 2010/04/09. DOI: 10.1097/PRS.0b013e3181df64fa.
27. Wood DJ and Gray B. Toward a Comprehensive Theory of Collaboration. *J App Behav Sci* 1991; 27: 139-162.
28. Devlin-Scherer R and Sardone NB. Collaboration as a Form of Professional Development: Improving Learning for Faculty and Students. *College Teaching* 2013; 61: 30-37. DOI: 10.1080/87567555.2012.714815.
29. Bronstein LR. A Model for Interdisciplinary Collaboration. *Social Work* 2003; 48: 297-306. DOI: 10.1093/sw/48.3.297.

30. Brass DJ, Galaskiewicz J, Greve HR, et al. Taking Stock of Networks and Organizations: A Multilevel Perspective. *Academy of Management Journal* 2004; 47: 795-817. DOI: 10.5465/20159624.

31. Chaskin RJ. Building Community Capacity. *Urban Affairs Review* 2016; 36: 291-323. DOI: 10.1177/10780870122184876.

32. Dyer JH and Nobeoka K. Creating and managing a high-performance knowledge-sharing network: the Toyota case. *Strategic Management Journal* 2000; 21: 345-367. DOI: 10.1002/(sici)1097-0266(200003)21:3<345::Aid-smj96>3.0.Co;2-n.

33. Gray B and Wood DJ. Collaborative Alliances Moving from Practice to Theory. *J App Behav Sci* 1991; 27: 3-22.

34. Miles RE, Snow CC and Pfeffer J. Organization-Environment: Concepts and Issues. *Industrial Relations* 1974; 13: 244-264. DOI: 10.1111/j.1468-232X.1974.tb00581.x.

35. Ring PS and van de Ven AH. Developmental Processes of Cooperative Interorganizational Relationships. *The Academy of Management Review* 1994; 19. DOI: 10.2307/258836.

36. Thomson AM and Perry JL. Collaboration Processes: Inside the Black Box. *Public Administration Review* 2006; 66: 20-32. DOI: 10.1111/j.1540-6210.2006.00663.x.

37. Colbry S, Hurwitz M and Adair R. Collaboration Theory. *J Lead Edu* 2014; 13: 63-75.

38. Patel H, Pettitt M and Wilson JR. Factors of collaborative working: a framework for a collaboration model. *Appl Ergon* 2012; 43: 1-26. 2011/05/28. DOI: 10.1016/j.apergo.2011.04.009.

39. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015; 4: 1. 2015/01/03. DOI: 10.1186/2046-4053-4-1.

40. McGowan J, Sampson M, Salzwedel DM, et al. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. *J Clin Epidemiol* 2016; 75: 40-46. 2016/03/24. DOI: 10.1016/j.jclinepi.2016.01.021.

41. Hooijmans CR, Rovers MM, de Vries RB, et al. SYRCLE's risk of bias tool for animal studies. *BMC Med Res Methodol* 2014; 14: 43. 2014/03/29. DOI: 10.1186/1471-2288-14-43.

42. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 Explanation and Elaboration: Updated guidelines for reporting parallel group randomised trials. *J Clin Epidemiol* 2010; 63: e1-37. 2010/03/30. DOI: 10.1016/j.jclinepi.2010.03.004.

43. Canada H. *E6(R2) Good Clinial Practice: Integrated Addendum to ICH E6(R1) - Guidance for Industry*. . 2016.

44. Frey BB, Lohmeier JH, Lee SW, et al. Measuring Collaboration Among Grant Partners. *American Journal of Evaluation* 2016; 27: 383-392. DOI: 10.1177/1098214006290356.

45. Popay J, Roberts H, Sowden A, et al. Guidance on the Conduct of Narrativ Synthesis in Systematic Reviews. 2006.

46. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009; 6: e1000097. 2009/07/22. DOI: 10.1371/journal.pmed.1000097.

47. Braun V and Clarke V. Using thematic analysis in psychology. *Qualitative Research in Psychology* 2006; 3: 77-101. DOI: 10.1191/1478088706qp063oa.

48. Srivastava A and Thompson BS. Framework Analysis: A Qualitative Methodology for Applied Policy Research. *Journal of Administration and Governance* 2009; 4.

49. Higgins J.P.T. and S. G. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. The Cochrane Collaboration, 2011.

50. Mondello S, Shear DA, Bramlett HM, et al. Insight into Pre-Clinical Models of Traumatic Brain Injury Using Circulating Brain Damage Biomarkers: Operation Brain Trauma Therapy. *J Neurotrauma* 2016; 33: 595-605. 2015/12/17. DOI: 10.1089/neu.2015.4132.

51. Kochanek PM, Bramlett HM, Dixon CE, et al. Approach to Modeling, Therapy Evaluation, Drug Selection, and Biomarker Assessments for a Multicenter Pre-Clinical Drug Screening Consortium

for Acute Therapies in Severe Traumatic Brain Injury: Operation Brain Trauma Therapy. *J Neurotrauma* 2016; 33: 513-522. 2015/10/07. DOI: 10.1089/neu.2015.4113.

52. Gill RG, Pagni PP, Kupfer T, et al. A Preclinical Consortium Approach for Assessing the Efficacy of Combined Anti-CD3 Plus IL-1 Blockade in Reversing New-Onset Autoimmune Diabetes in NOD Mice. *Diabetes* 2016; 65: 1310-1316. 2016/01/01. DOI: 10.2337/db15-0492.
53. Kochanek PM, Bramlett HM, Shear DA, et al. Synthesis of Findings, Current Investigations, and Future Directions: Operation Brain Trauma Therapy. *J Neurotrauma* 2016; 33: 606-614. 2015/12/17. DOI: 10.1089/neu.2015.4133.
54. Spoerke N, Zink K, Cho SD, et al. Lyophilized Plasma for Resuscitation in a Swine Model of Severe Injury. *Arch Surg* 2009; 144: 829-834.
55. Hagaman AK and Wutich A. How Many Interviews Are Enough to Identify Metathemes in Multisited and Cross-cultural Research? Another Perspective on Guest, Bunce, and Johnson's (2006) Landmark Study. *Field Methods* 2016; 29: 23-41. DOI: 10.1177/1525822x16640447.
56. Francis JJ, Johnston M, Robertson C, et al. What is an adequate sample size? Operationalising data saturation for theory-based interview studies. *Psychol Health* 2010; 25: 1229-1245. 2010/03/06. DOI: 10.1080/08870440903194015.
57. Sandelowski M. Sample size in qualitative research. *Research in Nursing & Health* 1995; 18: 179-183. DOI: 10.1002/nur.4770180211.
58. Strimbu K and Tavel JA. What are biomarkers? *Curr Opin HIV AIDS* 2010; 5: 463-466. 2010/10/28. DOI: 10.1097/COH.0b013e32833ed177.
59. Dechartres A, Boutron I, Trinquart L, et al. Single-center trials show larger treatment effects than multicenter trials: evidence from a meta-epidemiologic study. *Ann Intern Med* 2011; 155: 39-51. 2011/07/06. DOI: 10.7326/0003-4819-155-1-201107050-00006.
60. Lalu MM, Sullivan KJ, Mei SH, et al. Evaluating mesenchymal stem cell therapy for sepsis with preclinical meta-analyses prior to initiating a first-in-human trial. *Elife* 2016; 5 2016/11/22. DOI: 10.7554/eLife.17850.

61. McIntyre LA, Moher D, Fergusson DA, et al. Efficacy of Mesenchymal Stromal Cell Therapy for Acute Lung Injury in Preclinical Animal Models: A Systematic Review. *PLoS One* 2016; 11: e0147170. 2016/01/29. DOI: 10.1371/journal.pone.0147170.

62. Voelkl B, Vogt L, Sena ES, et al. Reproducibility of preclinical animal research improves with heterogeneity of study samples. *PLoS Biol* 2018; 16: e2003693. 2018/02/23. DOI: 10.1371/journal.pbio.2003693.

63. Baker D, Lidster K, Sottomayor A, et al. Two years later: journals are not yet enforcing the ARRIVE guidelines on reporting standards for pre-clinical animal studies. *PLoS Biol* 2014; 12: e1001756. 2014/01/11. DOI: 10.1371/journal.pbio.1001756.

64. Akison LK, Kuo J, Reid N, et al. Effect of Choline Supplementation on Neurological, Cognitive, and Behavioral Outcomes in Offspring Arising from Alcohol Exposure During Development: A Quantitative Systematic Review of Clinical and Preclinical Studies. *Alcohol Clin Exp Res* 2018; 42: 1591-1611. 2018/06/22. DOI: 10.1111/acer.13817.

65. Collins FS and Tabak LA. Policy: NIH plans to enhance reproducibility. *Nature* 2014; 505: 612-613. 2014/02/01. DOI: 10.1038/505612a.

66. McNutt M. Journals unite for reproducibility. *Science* 2014; 346: 679. 2014/11/11. DOI: 10.1126/science.aaa1724.

67. Neill US. Publish or perish, but at what cost? *J Clin Invest* 2008; 118: 2368. 2008/07/04. DOI: 10.1172/JCI36371.

68. Publish or perish. *Nature* 2010; 467: 252. 2010/09/17. DOI: 10.1038/467252a.

69. Thomson AM, Perry JL and Miller TK. Conceptualizing and Measuring Collaboration. *Journal of Public Administration Research and Theory* 2007; 19: 23-56. DOI: 10.1093/jopart/mum036.

70. Gravel K, Legare F and Graham ID. Barriers and facilitators to implementing shared decision-making in clinical practice: a systematic review of health professionals' perceptions. *Implement Sci* 2006; 1: 16. 2006/08/11. DOI: 10.1186/1748-5908-1-16.

71. Legare F, Ratte S, Gravel K, et al. Barriers and facilitators to implementing shared decision-making in clinical practice: update of a systematic review of health professionals' perceptions. *Patient Educ Couns* 2008; 73: 526-535. 2008/08/30. DOI: 10.1016/j.pec.2008.07.018.

72. Byrne MM, Tannenbaum SL, Gluck S, et al. Participation in cancer clinical trials: why are patients not participating? *Med Decis Making* 2014; 34: 116-126. 2013/07/31. DOI: 10.1177/0272989X13497264.

73. Gul RB and Ali PA. Clinical trials: the challenge of recruitment and retention of participants. *J Clin Nurs* 2010; 19: 227-233. 2010/05/27. DOI: 10.1111/j.1365-2702.2009.03041.x.

74. Johnson JA, Gong L, Whirl-Carrillo M, et al. Clinical Pharmacogenetics Implementation Consortium Guidelines for CYP2C9 and VKORC1 genotypes and warfarin dosing. *Clin Pharmacol Ther* 2011; 90: 625-629. 2011/09/09. DOI: 10.1038/clpt.2011.185.

75. Mahmud A, Zalay O, Springer A, et al. Barriers to participation in clinical trials: a physician survey. *Curr Oncol* 2018; 25: 119-125. 2018/05/03. DOI: 10.3747/co.25.3857.

76. Bakken S, Lantigua RA, Busacca LV, et al. Barriers, enablers, and incentives for research participation: a report from the Ambulatory Care Research Network (ACRN). *J Am Board Fam Med* 2009; 22: 436-445. 2009/07/10. DOI: 10.3122/jabfm.2009.04.090017.

77. Johnston S, Liddy C, Hogg W, et al. Barriers and facilitators to recruitment of physicians and practices for primary care health services research at one centre. *BMC Med Res Methodol* 2010; 10: 109. 2010/12/15. DOI: 10.1186/1471-2288-10-109.

78. Bafeta A, Dechartres A, Trinquart L, et al. Impact of single centre status on estimates of intervention effects in trials with continuous outcomes: meta-epidemiological study. *BMJ* 2012; 344: e813. 2012/02/16. DOI: 10.1136/bmj.e813.

79. Unverzagt S, Prondzinsky R and Peinemann F. Single-center trials tend to provide larger treatment effects than multicenter trials: a systematic review. *J Clin Epidemiol* 2013; 66: 1271-1280. 2013/08/27. DOI: 10.1016/j.jclinepi.2013.05.016.

80. Kilkenny C, Browne WJ, Cuthill IC, et al. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *Osteoarthritis Cartilage* 2012; 20: 256-260. 2012/03/20. DOI: 10.1016/j.joca.2012.02.010.

81. Ghaffar A, Langlois EV, Rasanathan K, et al. Strengthening health systems through embedded research. *Bull World Health Organ* 2017; 95: 87. 2017/03/03. DOI: 10.2471/BLT.16.189126.

82. Lasker RD, Weiss ES and Miller R. Partnership Synergy: A Practical Framework for Studying and Strengthening the Collaborative Advantage. *The Milbank Quarterly* 2003; 79: 179-205. DOI: 10.1111/1468-0009.00203.

83. Abbas I, Rovira J, Casanovas J, et al. Optimal design of clinical trials with computer simulation based on results of earlier trials, illustrated with a lipodystrophy trial in HIV patients. *J Biomed Inform* 2008; 41: 1053-1061. 2008/06/07. DOI: 10.1016/j.jbi.2008.04.008.

84. Moher D, Schulz KF, Simera I, et al. Guidance for developers of health research reporting guidelines. *PLoS Med* 2010; 7: e1000217. 2010/02/20. DOI: 10.1371/journal.pmed.1000217.

85. Bailey KR. Detecting Fabrication of Data in a Multicenter Collaborative Animal Study. *Controlled Clinical Trials* 1991; 12: 741-752.

86. Mountney A, Bramlett HM, Dixon CE, et al. Simvastatin Treatment in Traumatic Brain Injury: Operation Brain Trauma Therapy. *J Neurotrauma* 2016; 33: 567-580. 2015/11/07. DOI: 10.1089/neu.2015.4130.

87. Bramlett HM, Dietrich WD, Dixon CE, et al. Erythropoietin Treatment in Traumatic Brain Injury: Operation Brain Trauma Therapy. *J Neurotrauma* 2016; 33: 538-552. 2015/12/17. DOI: 10.1089/neu.2015.4116.

88. Browning M, Shear DA, Bramlett HM, et al. Levetiracetam Treatment in Traumatic Brain Injury: Operation Brain Trauma Therapy. *J Neurotrauma* 2016; 33: 581-594. 2015/12/17. DOI: 10.1089/neu.2015.4131.

89. Dixon CE, Bramlett HM, Dietrich WD, et al. Cyclosporine Treatment in Traumatic Brain Injury: Operation Brain Trauma Therapy. *J Neurotrauma* 2016; 33: 553-566. 2015/12/17. DOI: 10.1089/neu.2015.4122.

90. Shear DA, Dixon CE, Bramlett HM, et al. Nicotinamide Treatment in Traumatic Brain Injury: Operation Brain Trauma Therapy. *J Neurotrauma* 2016; 33: 523-537. 2015/12/17. DOI: 10.1089/neu.2015.4115.

Appendices

Appendix 1: Search strategies

A – original search terms in Medline, B – update search terms in MEDLINE, C – original search terms in Embase, D – updated search terms in Embase

A

Database: Ovid MEDLINE(R) ALL <1946 to May 03, 2019>
Search Strategy: Up to April 24, 2018
---
1 Drug Evaluation, Preclinical/ (48152)
2 exp models, animal/ (534660)
3 Animals, Laboratory/ or exp *animals, laboratory/ (23855)
4 (preclinic* or pre clinic*).ti. (15707)
5 (animal* and model*).ti. (23378)
6 or/1-5 (613079)
7 multicenter study.pt. (249599)
8 ((cross or across) adj2 (lab or labs or laborator*)).tw. (1882)
9 ((collabor* or cooperativ* or multisite or multi site or global) adj2 (stud* or trial* or experiment*)).tw. (24676)
10 ((multi* institut* or multiple* site*) adj5 (trial* or stud* or experiment*)).tw. (4729)
11 ((multicent* or multi cent*) and (trial* or stud* or experiment*)).tw,kw. (140854)
12 (consortia* or consortium).tw,kw. (24678)
13 or/7-12 (352955)
14 6 and 13 (1218)
15 ((multicent* or multi cent*) and (preclinic* or pre clinic* or experiment*)).ti. (86)
16 14 or 15 (1266)
17 ("20180425" or "20180426" or "20180427" or "20180428" or "20180429" or 2018043* or 201805* or 201806* or 201807* or 201808* or 201809* or 20181* or 2019*).dt. (1340703)
18 16 not 17 (1211)
19 remove duplicates from 18 (1211)

B

Database: Ovid MEDLINE(R) ALL <1946 to May 03, 2019>
Search Strategy: April 24, 2018
---
1 Drug Evaluation, Preclinical/ (48152)
2 exp models, animal/ (534660)
3 Animals, Laboratory/ or exp *animals, laboratory/ (23855)
4 (preclinic* or pre clinic*).ti. (15707)
5 (animal* and model*).ti. (23378)
6 or/1-5 (613079)
7 ((multi* institut* or multiple* site*) adj5 (trial* or stud* or experiment*)).tw. (4729)
8 6 and 7 (20)
9 ("20180425" or "20180426" or "20180427" or "20180428" or "20180429" or 2018043* or 201805* or 201806* or 201807* or 201808* or 201809* or 20181* or 2019*).dt. (1340703)
10 8 and 9 (0) April 2018 - present
11 8 not 10 (20) Inception to April 2018

C

Database: Embase Classic+Embase <1947 to 2019 May 03>
Search Strategy: Up to April 24, 2018

---

1 animal experiment/ (2361114)
2 experimental animal/ (29269)
3 animal model/ (1212440)
4 (preclinic* or pre clinic*).ti. (24531)
5 (animal* and model*).ti. (31203)
6 or/1-5 (2595300)
7 multicenter study/ (214255)
8 ((cross or across) adj2 (lab or labs or laborator*)).tw. (2592)
9 ((collabor* or cooperativ* or multisite or multi site or global) adj2 (stud* or trial* or experiment*)).tw. (33026)
10 ((multicent* or multi cent*) adj5 (trial* or stud* or experiment*)).tw. (193832)
11 (multi* institut* adj5 (trial* or stud* or experiment*)).tw. (6990)
12 (consortia* or consortium).tw. (40093)
13 or/7-12 (387072)
14 6 and 13 (5779)
15 ((multicent* or multi cent*) and (preclinic* or pre clinic* or experiment*)).ti. (117)
16 14 or 15 (5845)
17 conference abstract.pt. (3396626)
18 conference review.pt. (11174)
19 16 not (17 or 18) (2619)
20 ("20180425" or "20180426" or "20180427" or "20180428" or "20180429" or 2018043* or 201805* or 201806* or 201807* or 201808* or 201809* or 20181* or 2019*).dc. (1917329)
21 19 not 20 (2220)

D

Database: Embase Classic+Embase <1947 to 2019 May 03>
Search Strategy: April 24, 2018

---

1 animal experiment/ (2361114)
2 experimental animal/ (29269)
3 animal model/ (1212440)
4 (preclinic* or pre clinic*).ti. (24531)
5 (animal* and model*).ti. (31203)
6 or/1-5 (2595300)
7 (multi* institut* adj5 (trial* or stud* or experiment*)).tw. (6990)
8 6 and 7 (71)
9 (conference abstract or conference review).pt. (3407800)
10 8 not 9 (31)
11 ("20180425" or "20180426" or "20180427" or "20180428" or "20180429" or 2018043* or 201805* or 201806* or 201807* or 201808* or 201809* or 20181* or 2019*).dc. (1917329)
12 10 and 11 (7) April 2018 - present
13 10 not 12 (24) Inception to April 2018

Appendix 2: PRESS review of search strategy

***PRESS Guideline* — Search Submission & Peer Review Assessment**

SEARCH SUBMISSION: THIS SECTION TO BE FILLED IN BY THE SEARCHER

Searcher's Name: Risa Shorr Email: rshorr@toh.ca

Date Submitted: 2018-01-172018-01-17

Systematic Review Title: Preclinical Multicenter Studies

This search strategy is: My PRIMARY (core) database strategy — First time submitting a strategy for search

Database (ie, Medline, Cinahl): Medline

Interface (Ovid, Ebsco): Ovid

Research Question. Describe the purpose of the search: To summarise the literature on preclinical multicenter studies

PICO Format (Outline the PICOs for your question – ie. Patient, Intervention, Comparison, Outcome and Study Design – as applicable)

P	Preclinical studies
I	Multicenter
C	None
O	None

Was a search filter applied? No

Other notes or comments you feel would be useful for the peer reviewer? Here are some target articles. All are captured with the search strategy. Searching all animal groups and study type is too huge (~10000 refs in Medline) and then not all studies are indexed as multicenter.

1. **A cross-laboratory preclinical study on the effectiveness of interleukin-1 receptor antagonist in stroke**.
Maysami S; Wong R; Pradillo JM; Denes A; Dhungana H; Malm T; Koistinaho J; Orset C; Rahman M; Rubio M; Schwaninger M; Vivien D; Bath PM; Rothwell NJ; Allan SM.
Journal of Cerebral Blood Flow & Metabolism. 36(3):596-605, 2016 Mar.
[Journal Article. Research Support, Non-U.S. Gov't]
UI: 26661169

2. **Results of a preclinical randomized controlled multicenter trial (pRCT): Anti-CD49d treatment for acute brain ischemia.**
Llovera G; Hofmann K; Roth S; Salas-Perdomo A; Ferrer-Ferrer M; Perego C; Zanier ER; Mamrak U; Rex A; Party H; Agin V; Fauchon C; Orset C; Haelewyn B; De Simoni MG; Dirnagl U; Grittner U; Planas AM; Plesnila N; Vivien D; Liesz A.
Science Translational Medicine. 7(299):299ra121, 2015 Aug 05.
[Journal Article. Multicenter Study. Research Support, Non-U.S. Gov't]
UI: 26246166

3. **The NHLBI-sponsored Consortium for preclinicAl assESsment of cARdioprotective therapies (CAESAR): a new paradigm for rigorous, accurate, and reproducible evaluation of putative infarct-sparing interventions in mice, rabbits, and pigs.**

Jones SP; Tang XL; Guo Y; Steenbergen C; Lefer DJ; Kukreja RC; Kong M; Li Q; Bhushan S; Zhu X; Du J; Nong Y; Stowers HL; Kondo K; Hunt GN; Goodchild TT; Orr A; Chang CC; Ockaili R; Salloum FN; Bolli R.
Circulation Research. 116(4):572-86, 2015 Feb 13.
[Journal Article. Multicenter Study. Research Support, N.I.H., Extramural]
UI: 25499773

4. **Different data from different labs: lessons from studies of gene-environment interaction. [Review] [83 refs]**
Wahlsten D; Metten P; Phillips TJ; Boehm SL 2nd; Burkhart-Kasch S; Dorow J; Doerksen S; Downing C; Fogarty J; Rodd-Henricks K; Hen R; McKinnon CS; Merrill CM; Nolte C; Schalomon M; Schlumbohm JP; Sibert JR; Wenger CD; Dudek BC; Crabbe JC.
Journal of Neurobiology. 54(1):283-311, 2003 Jan.
[Comparative Study. Journal Article. Research Support, Non-U.S. Gov't. Research Support, U.S. Gov't, Non-P.H.S.. Research Support, U.S. Gov't, P.H.S.. Review]
UI: 12486710

5. **Genetics of mouse behavior: interactions with laboratory environment**.
Crabbe JC; Wahlsten D; Dudek BC.
Science. 284(5420):1670-2, 1999 Jun 04.
[Comparative Study. Journal Article. Research Support, Non-U.S. Gov't. Research Support, U.S. Gov't, Non-P.H.S.. Research Support, U.S. Gov't, P.H.S.]
UI: 10356397

6. **Animal models for protecting ischemic myocardium: results of the NHLBI Cooperative Study. Comparison of unconscious and conscious dog models.**
Reimer KA; Jennings RB; Cobb FR; Murdock RH; Greenfield JC Jr; Becker LC; Bulkley BH; Hutchins GM; Schwartz RP Jr; Bailey KR; et al.
Circulation Research. 56(5):651-65, 1985 May.
[Comparative Study. Journal Article. Research Support, U.S. Gov't, P.H.S.]
UI: 3838923

Please copy and paste your search strategy here, exactly as run, including the number of hits per line.

Database: Ovid MEDLINE(R) ALL <1946 to January 16, 2018>
Search Strategy:
---
1 Drug Evaluation, Preclinical/ (49649)
2 exp models, animal/ (552406)
3 Animals, Laboratory/ or exp *animals, laboratory/ (25944)
4 (preclinic* or pre clinic*).ti. (15572)
5 or/1-4 (626909)
6 multicenter study.pt. (261962)
7 ((cross or across) adj2 (lab or labs or laborator*)).tw. (1811)
8 ((collabor* or cooperativ* or multisite or multi site or global) adj2 (stud* or trial* or experiment*)).tw. (25511)
9 ((multicent* or multi cent*) and (trial* or stud* or experiment*)).tw,kw. (143765)

10 6 or 7 or 8 or 9 (340472)
11 5 and 10 (927)
12 ((multicent* or multi cent*) and (preclinic* or pre clinic* or experiment*)).ti. (77)
13 11 or 12 (977)

PEER REVIEW ASSESSMENT: THIS SECTION TO BE FILLED IN BY THE REVIEWER

Reviewer: Sascha Davis adavis@toh.ca Date Completed:2018-01-242018-01-24

Translation: No Revision
Boolean and Proximity Operators: No Revision
Subject Headings: No Revision
Text Word Searching: No Revision
Spelling, Syntax and Line Numbers: No Revision
Limits and Filters: No Revision
Overall Evaluation: No Revision
Additional Comments:

- Could you do "animal model*".ti or is that too big? Or even "animal adj2 model*".ti?
- I'm not sure if more wording could be added for the concept of cross-laboratory? – "different labs" or "simultaneous* adj3 laborator*) or ("two laborator* or "three laborator* or "four laborator*")?

Appendix 3: Multicenter reporting checklist with item domain and source(s)

Domain	#	Item Description	Source(s)
Intro/abstract	1	Identification as a multicenter study in title	CONSORT
	2	Abstract states number of participating centers	CONSORT
Standards	3	Community based reporting guidelines listed	NIH
	4	Names of each participating center listed	GCP E6(R2)
	5	List roles of participating centers (central coordinating center, experimental site)	GCP E6(R2)
	6	No changes, or if applicable major changes to study protocol after commencement are documented	CONSORT
Replicates (biological vs. technical)	7	Results substantiated by repetition under a range of conditions at each site	NIH, CONSORT
	8	Number of subjects per outcome	NIH, CONSORT
	9	Number of measurements per subject for one experimental outcome stated	NIH, CONSORT
	10	Number of subjects per center	GCP E6(R2)
Statistics	11	List of the total number of subjects used in each experimental group	NIH, CONSORT
	12	List of all statistical tests used	NIH, CONSORT
	13	Definition of the measure of central tendency	NIH
	14	Definition of the measure of dispersion	NIH
Randomization	15	Random group assignment reported	NIH, CONSORT
	16	Description of the method of random group assignment	NIH, CONSORT
Blinding	17	Experimenters blinded to group allocation during conduct of the experiment	NIH, CONSORT
	18	Experimenters blinded to group allocation during result assessment	NIH, CONSORT
Sample Size Estimation	19	Description of an *a priori* primary outcome	CONSORT
	20	Sample size computed during study design	NIH, CONSORT
	21	Description of the method of sample size determination	NIH, CONSORT
Inclusion and Exclusion Criteria	22	Total number of animals for the experiment reported	GCP E6(R2)
	23	Description of the criteria used for the exclusion of any data or subjects	NIH, CONSORT
	24	List losses and exclusions of animals at the end of experiment	CONSORT
	25	All outcomes described, or description of any outcomes measured but not reported in results	NIH, CONSORT
	26	Previous or pilot/preliminary studies performed and listed	NIH
	27	Results were significant, or if not, null or negative outcomes included in the results	NIH
Discussion	28	Limitations of the study are documented	CONSORT
	29	Discrepancies in results across centers expected or absent, or if not, they discussed	CONSORT

CONSORT: Consolidated Standards of Reporting Trials
NIH: National Institutes of Health
GCP E6(R2): Guideline for Good Clinical Practice, 2016 integrated addendum

Appendix 4: Degree of collaboration assessment criteria for domains

Domain	DOC Reported			
	Low = 1	Medium = 2	High = 3	Unclear = 0
Development	Not all centers involved in protocol development; no pilot study	Protocol developed *a priori* by all centers; no pilot test	Protocol developed *a priori* by all centers with pilot study	Unclear which centers were involved in protocol development; no pilot study
Execution	Shared outcomes and design, different methods of data collection and analysis between centers	Shared protocol with minor differences in methods	Shared protocol and analysis for all participating centers	Unclear if protocol and analysis was shared across all centers
Coordination	No coordinating committee/center	Committee responsible for coordination, no independent coordinating center	Independent center designated to study coordination	Unclear how and by who the study was coordinated

Appendix 5: PRISMA Checklist

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	**1**	**Identify the report as a systematic review, meta-analysis, or both.**	**NA**
ABSTRACT			
Structured summary	**2**	**Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.**	**ii**
INTRODUCTION			
Rationale	**3**	**Describe the rationale for the review in the context of what is already known.**	**4**
Objectives	**4**	**Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).**	**4-5**
METHODS			
Protocol and registration	**5**	**Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.**	**12**
Eligibility criteria	**6**	**Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.**	**12-13**
Information sources	**7**	**Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.**	**13-14**
Search	**8**	**Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.**	**101-102**
Study selection	**9**	**State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).**	**14-15**
Data collection process	**10**	**Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.**	**15**
Data items	**11**	**List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.**	**15**
Risk of bias in individual studies	**12**	**Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.**	**15-16**
Summary measures	**13**	**State the principal summary measures (e.g., risk ratio, difference in means).**	**NA**
Synthesis of results	**14**	**Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.**	**NA**
Risk of bias across studies	**15**	**Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).**	**NA**
Additional analyses	**16**	**Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.**	**NA**
RESULTS			

Study selection	**17**	**Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.**	**23-24**
Study characteristics	**18**	**For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.**	**25-26**
Risk of bias within studies	**19**	**Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).**	**28-29**
Results of individual studies	**20**	**For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.**	**27-28**
Synthesis of results	**21**	**Present results of each meta-analysis done, including confidence intervals and measures of consistency.**	**NA**
Risk of bias across studies	**22**	**Present results of any assessment of risk of bias across studies (see Item 15).**	**NA**
Additional analysis	**23**	**Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).**	**NA**
DISCUSSION			
Summary of evidence	**24**	**Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).**	**66-69**
Limitations	**25**	**Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).**	**69**
Conclusions	**26**	**Provide a general interpretation of the results in the context of other evidence, and implications for future research.**	**87-90**
FUNDING			
Funding	**27**	**Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.**	**NA**

Appendix 6: Interview guide

Exploring collaboration in preclinical multicenter studies: Interview Guide for researchers who have conducted a multicenter study

Explanation

Thank you for agreeing to speak with me today about your experience and perspective on preclinical multicenter studies. The objectives of this study are to identify experiences, opinions, and potential barriers and facilitators to conducting such studies. Additionally, I am also interested in comparing these experiences of multicenter studies to those in single-centered studies. The interview should take approximately 30 minutes and will be audio-recorded to ensure that all key points are accurately documented. Just so you are aware, I am a Master's student in Health Systems therefore I am not an expert in preclinical research - so I may ask you to clarify and explain throughout the course of our discussion. Is this alright with you?

Any identifying information (for example, your name or names of other individuals) that you use in the course of our discussion will be removed from the interview transcripts. If you wish to end the interview before I have asked all of the questions or if you wish to withdraw from the study, you are free to do so, just let me know. It is important to note that there are no right or wrong answers to the questions.

Do you have any questions before we start?

As you know, you were contacted for an interview because you participated in the multicenter study: "***Study name*"**, published in "***journal***". For simplicity, I'll refer to this as 'this study' or 'you study' for the remainder of our discussion.

Questions

1. Can you briefly tell me what you did in this preclinical multicenter study?
   - Can you tell me about the study protocol and the experiments?

2. What can you tell me about your overall experience working in a preclinical multicenter study?
   - Did you enjoy collaborating in this preclinical multicenter study?
   - What did you enjoy about collaborating?
   - What didn't you enjoy about collaborating?

3. What was your specific role in the multicenter study?

4. Can you tell me about the development of the study protocol?
   - Who was involved in its development?
   - How was the protocol developed (Pilot tests, based off the literature)?
   - How long did the process take?
   - What were some of the challenges with developing the protocol across sites?

5. In general, what were your thoughts of preclinical multicenter studies *before* conducting/being involved with one?

PROMPTS:
- Did you believe they were superior to single-centered studies?
- Did your opinions change after the completion of the study?

6. Prior to conducting this preclinical multicenter study we are discussing today, what were your thoughts and expectations about collaborating in this study?

PROMPTS:

- Did you believe that the study would run smoothly, or did you expect a few hiccups?
- Did you expect that the study would be a positive or negative experience?
- Did you feel confident that you could collaborate in this study?
- What made you, or didn't make you confident

7. Why did you decide to collaborate on this preclinical multicenter study (i.e. use a preclinical multicenter study design)?

PROMPTS:

- What influenced your decision to collaborate? OR
- Was there anybody that influenced you to collaborate?

8. How was the collaboration between centers in the study initiated?

PROMPTS:

- Did any members have previously established relationships?
- Where did the idea to conduct a multicenter study originate from (please elaborate)?

9. Did you experience any challenges to collaboration in the multicenter study?

PROMPTS:

- Were they foreseeable?
- Were there strategies in place to mitigate them? If yes, please elaborate.
- Where they overcome? If yes, how so?

10. Were there any benefits or rewards to collaboration in the multicenter study?
    - Was there anything about the study that made it work so well (with no challenges)?

11. Tell me about communication with the team during the multicenter study.

PROMPTS:

- Who did you communicate with? And how frequently?
- How did you, and other participating members communicate? (i.e. In-person, email, conferences, phone call, text) throughout the course of the study?
- Did you communicate with members outside of the center you worked at/with originally?

12. Were there any elements of the project that were essential to the study's execution?

PROMPTS:

- If yes, please explain.
- Was there an individual or individuals that had a leadership role in the study?
- Were there any shared rules or norms?

13. What things do you think make a team ready to take on a multicenter study?

PROMPTS:

- Are there any skills, resources or measures needed to succeed in using this method?

14. Can you tell me about the writing (drafting and editing) of the results, and final manuscript?

PROMPTS:

- Was everyone who was involved in the protocol development and execution involved in writing the manuscript?

15. Knowing what you know about conducting/being involved in a multicenter study, if you were to do it over, would there be anything that you would do differently?

PROMPTS:

- How likely would you be to conduct/be involved in another multicenter study (if it was your choice)?

16. Overall, do you believe that this preclinical multicenter was worthwhile?

PROMPTS:

- Was it worth the challenges and work involved?
- Was is a worthwhile experience for you personally?
- Do you believe it was valuable to the field of science?

Experience with single-centered-studies compared to multicenter studies

Another aim of my study is to compare the experiences, views, and barriers and facilitators between single-centered studies and multicentered studies. Can you state whether you have you previously been involved in a single-center preclinical study? If yes:

1. Tell me about you experiences with single-centered studies as compared to multicenter studies.
2. Are there any advantages to a single-centered design over a multicenter design?

PROMPTS:

- If yes, what are they?

3. What are some advantages of multicentered over single-centered design?

Background of respondent

Next, I have a few demographic questions. Your specific answers will not be tied back to you, but please only answer if you feel comfortable.

1. What area or field of preclinical research do you study, and have studied in the past?
2. What is the title of your current position?

PROMPT (if not clear): What are the responsibilities of this position?

3. (For PIs) Around how many years have you been an independent investigator with your own lab?

(For Co-Is) Around how many years have you been involved with preclinical research (including graduate school)?

4. For demographic purposes, could you tell me what gender you identify with?

Those are all the questions I have for you. Is there anything else you would like to expand upon?

We are also looking to recruit more interview participants through snowball sampling methods. Would you mind connecting me with other site PIs with whom you collaborated with in this study? If it is alright with you, I can send you a follow-up reminder email, and then you will be able to send me their contact information.

Appendix 7: Code book with definitions

Code	Definition
Stakeholder autonomy	Collaborators retaining/relinquishing decision making powers. They collaborate by their own free will or they were forced to collaborate from external pressure.
Interactive process	The change-oriented relationship. The process of the collaboration develops interactively between collaborators. (Natural shifts in structure of collaboration/evolution, collaborator working together)
Shared rules, norms and structures	Any rules, norms and structures that have been developed explicitly by the collaborators during the collaborative process.
The action	The goal of objective of the collaboration. The collaborators demonstrate this by expressing priorities and end goals. (Reasons for collaboration/what they are trying to achieve)
Domain orientation	Collaborators orienting their processes, decisions, actions and attention to the issues of the problem domain that brought them together (action planning, decision making).
Convenor	An individual who initiates or facilitates the formation of a collaboration. Identifies and brings collaborators together.
Roles and Identity	Any specific roles, identities or personal qualities within the collaboration that may affect the collaborative process.
Benefits	Any real or perceived positive consequence of the collaboration.
Costs	Any real of perceived negative consequence of the collaboration.
Barriers	Any real or perceived barrier, difficulty, hurdle or problem that may have impeded the collaboration from reaching its goal.
Enablers	Any real or perceived element or aspect that enabled or facilitated the collaboration in reaching its goal.
Expectations	Expectations, predictions or anticipated results of the collaboration
Attitudes	Any expression of the thoughts and feelings towards the preclinical multicenter collaboration
Single-center comparison	Any comparison of views, experiences, goals or structures between single-center and multicenter studies
Clinical trial comparison	Any comparison of views, experiences, goals or structures between clinical trials and multicenter studies
Study description	Any discussion around the study details/logistics.
Control	Discussion around having or not having control (personal or organizational) during the course of the study
The multicenter design	Any discussion of the general purpose, importance, use and appropriateness of a multicenter study design in preclinical research

Appendix 8: Sub-code book with definitions

Code	Sub-Code	Sub-Code Definition
Roles and Identity	Professional role	Discussion about the professional role of the collaborator or other members of the collaboration during the study
	Personal role	Discussion about the personal role of the collaborator or other members of the collaboration during the study
	Past experiences	Past experience of the collaborators that influenced the conduct of the study
Benefits	Learning	Mention of the study as a learning experience to the collaborators
	Reaching goals	Mention of reaching the goals of the study
	Rewards	Any additional rewards achieved from being involved in the collaboration
Costs	Stress	Mention of stress to the collaborators due to the multicenter study
	Monetary	Personal monetary/transaction costs on the collaborators themselves due to the multicenter study
	Regret	Mention of regret in decision to be involved in multicenter study, or regret of specific aspect of collaboration
Barriers	Financial/resources	Any financial/resource/material aspects that hinder the conduct of the study (Ex. lack of or not enough)
	Environment	Mention of the lab or research environment as a barrier to the conduct of the study
	Culture/climate	Mention of the culture/climate of the scientific community as a barrier to conducting the study
	Critical incidents	Any specific critical incidents that hindered the conduct of the study
	Social	Any personal or social barrier that hindered the conduct of the study
Enablers	Skills	Any skills of the collaborator that facilitate the conduct of the study
	Personal/social	Personal or social aspects present in the coloration that facilitated the conduct of the study
	Financial/resource	Finances and/or resources that facilitated the conduct of the study
	Environment	Environmental culture or context that was supportive to the conduct of the study
	Trust/transparency	Trust and transparency between member of the collaboration to facilitate the conduct of the study
Attitudes	Positive attitude	Mention of positive attitudes, perspectives or views towards the multicenter collaboration
	Negative attitude	Mention of negative attitudes, perspectives or views towards the multicenter collaboration
Single center comparison	Environmental complexity	Comparison or mention of the difference in environmental complexity between single-center and multicenter studies
Study description	Study length	Description of the time it took from study conception, conduct and dissemination of the results
	Funding	Description of how the study was funded and where funding came from
	Site selection	Description of how and why the study sites were selected
	Logistics	Mention of logistical details of the study (daily business, shipping samples)
	Dissemination	Description of the dissemination of the study results (publications, writing process)
The multicenter design	Appropriateness	Appropriateness of the multicenter design for the research question that is being answered
	Contributions	How the multicenter study, or the design is making a contribution to the scientific community/field of science
	Purpose of design	Discussion of the reason for why a multicenter design is used in preclinical research
	Future research	Hypothetical or real application of multicenter design to future research or hypothetical changes to complete study

Appendix 9: Study Overviews

1. In a study by Reimer et al. (1985) [25], three independent laboratories collaborated to develop models to test potential ischemic myocardium protection therapies, using two standardized, well-characterized canine models of myocardial infarction. Using the two different dog models (conscious model of coronary occlusion, and unconscious model of 3-hour ischemia in open-chest), the researchers tested the effects of verapamil and ibuprofen (therapies) on infarct size. The pooled results from all three centers demonstrated that neither drug limited infarct size in either model. It was later published that the participating laboratories discovered through statistical and hard evidence that a forth participating lab initially involved in the study had fraudulent data, in the sense that data had been completely fabricated by the lead researcher at one of the centers [85]. The data from this lab was not included in the multicenter study paper. The detection of the fraudulent data would not have been possible if not for the design of a multicenter study. The fraud was detected by the large discrepancies of study outcome data between to offending center and the other centers involved in the study. It was later confirmed by the coordinating as they performed own investigation into the lead scientist of the offending center.

2. Crabbe et al. (1999) [24] performed a large study across 3 laboratories. The main objective was to test the behavioural variability in mice of different genetic strains, sexes, and laboratory environments. The evaluation was done with identical testing apparatus and protocol across all three labs. The potentially clinically relevant portion of this study was an assessment of cocaine's effect on behavior (i.e. locomotor activity). The study found that cocaine effects on locomotor activity had a strong relationship with genetic

differences on the laboratory giving the tests but was negligible for sex differences and source of mice (i.e. shipped from a supplier or were bred locally).

3. Alam et al. (2009) (cite) conducted a three-phase severe traumatic injury protocol to model trauma-induced coagulopathy, acidosis, and hypothermia on Yorkshire swine across three experimental centers. Animals were treated with 4 different blood products: fresh whole blood (FWB), hetastarch, fresh frozen plasma/packed RBCs (FFP: PRBC), and FFP, to determine which, if any, were effective in reversing trauma-associated coagulopathy. Treatment with FFP and FFP: PRBC corrected the coagulopathy aa effectively as FWB, whereas hetastarch worsened coagulopathy.

4. Spoerke et al. (2009) [54] tested whether lyophilized plasma (LP) is as safe and effective as fresh frozen plasma (FFP) for resuscitation after severe trauma. They used a swine model of severe injury across animal laboratories of two level I trauma centers, to test the lyophilized plasma for factor levels and clotting activity before lyophilization and after reconstitution. The swine model was developed and performed at one of the centers and was learned and performed at a second center to test for reproducibility. They found that LP decreased clotting factor activity and was equal to FFP in terms of efficacy.

5. Jones et al. (2015) [20] aimed to develop a multicenter, randomized controlled clinical-like infrastructure for preclinical evaluation of cardioprotective therapies using mice, rabbit and pig models. The researchers established the Consortium for preclinical assessment of cardioprotective therapies – called CAESAR, to test the ability of ischemic preconditioning (IPC) to reduce infarct size of a myocardial infarction. IPC involved short episodes of blood restriction to the heart – which is an experimental technique for producing resistance to blood loss. It has demonstrated to activate the largest number of

protective pathways and is currently the most reproducible cardioprotection intervention to date. Six centers (2 centers/animal model) tested the therapy in the three animal models with shared protocols, and found the results were similar across centers and that IPC significantly reduced infarct size in all three species.

6. Llovera and colleagues (2015) [22] performed a preclinical randomized controlled multicenter trial (pRCT) to test the potential of Anti-CD49d antibodies as treatment from acute brain ischemia. These antibodies have shown promise as a form of therapy in individual laboratories by inhibiting the migration of leukocytes into the brain following acute brain lesion. Leukocyte invasion is known as brain inflammation and is a key mechanism that mediates secondary neuronal injury after a stroke. Six independent European research centers centrally coordinated a study to test the antibody using mouse models of stroke, and the pooled results demonstrated that the antibody significantly reduced leukocyte invasion after mechanically induced stroke (distal occlusion of the middle cerebral artery). They found that the treatment did not limit infarct size and concluded that this therapy should not be evaluated further in clinical trials.

7. Maysami and colleagues (2015) [23] conducted a cross-laboratory study in five centers (4 research, 1 coordinating) to test and interleukin receptor antagonist as a drug therapy for stroke. The antagonist drug is a naturally occurring drug that has been reported as beneficial for stroke recovery, as it inhibits the neuronal binding of the cytokine interleukin – a key mediator in neuronal injury (inflammation). The coordinating center developed and distributed the standard operating procedure to all centers, where mice were used as the model and ischaemia was induced both by permanent and transient occlusion. Drug effects on stroke outcome was evaluated by various means: lesion

volume, oedema, neurological deficit scoring and post-treatment mortality. The results across all centers strongly support the therapeutic potential of the cytokine receptor antagonist in experimental stroke.

8. Five separate studies [86-90] that were coordinated by Operation Brain Trauma Therapy (OBTT) consortium [50, 51, 53]. Three independent centers collaborated to screen 5 different drugs that had been proposed for acute therapies in severe traumatic brain injury (TBI). The consortium was supported by the United States Army and had an overall approach of testing promising therapies in three-well established models of TBI in rats with a rigorous design. The end goal of the consortium was to test the 5 initial therapies in rats prior to considering further testing in a swine model of TBI. Based on the results, four of the five drugs preformed below or well below what was expected based on the previously published literature. It was reported that only levetiracetam would advance to testing in the swine model, and that an additional three drugs were being tested by the consortium as well.

9. Gill et al. (2016) [52] assessed the efficacy of combined anti-CD3 plus interleukin-1 blockade to reverse new-onset autoimmune diabetes in non-obese diabetic (NOD) mice. Their consortium was established by the National Institutes of Health Immune Tolerance Network and the Juvenile Diabetes Research Foundation. Four academic centers shared models and operating procedures, in nine different treatment groups. They found that the combined antibody treatment did not show reversal of diabetes across all sites. They did however conclude that intercenter reproducibility is possible with the NOD mouse model of diabetes.

Appendix 10: Statements of future recommendations

Author, Year	Recommendation statements
Reimer, 1985	Nothing reported
Crabbe, 1999	*Relatively small genetic effects should first be replicated locally before drawing conclusions... genotypes should be tested in multiple labs and evaluated with multiple tests of a single behavioral domain*
Alam, 2009	*Based upon the findings of the current study that demonstrated the impressive hemostatic properties of plasma, we have proceeded to successfully develop and test (in the same model) lyophilized FDP.*
Spoerke, 2009	*The species-specific differences in factor activities will require ongoing investigation to ensure full safety and efficacy. Our future investigations will include a comprehensive evaluation of the effects of the lyophilization process on coagulation properties of the LP.*
Jones, 2015	*other investigators can adopt the protocols [for measuring infarct size in mice, rabbits, and pigs in a manner that is rigorous, accurate, and reproducible] in their own laboratories.*
Llovera, 2015	*future clinical trials testing immunotherapeutic drugs for stroke will need to ensure that the included study population feature a substantial neuroinflammatory reaction to the brain injury*
Maysami, 2015	*interleukin 1 receptor antagonist should be evaluated in more extensive clinical stroke trials*
Bramlett, 2016	*Although we cannot rule out the possibility that other doses or more prolonged treatment could show different effects, the lack of efficacy of EPO reduced enthusiasm for its further investigation in OBTT.*
Browning, 2016	*…need for OBTT to study LEV further. This includes studies of dose response, therapeutic window, mechanism, and testing in our large animal FPI model in micropigs… consider a randomized controlled trial examining early administration in patients*
Dixon, 2016	*Our findings reduce enthusiasm for further investigation of this therapy in OBTT and suggest that if this strategy is to be pursued further, alternative CsA analogs with reduced toxicity should be used.*
Gill, 2016	*…pause in proceeding with clinical trials without further preclinical testing.*
Mountney, 2016	*the current findings do not support the beneficial effects of simvastatin… it will not be further pursued by OBTT.*
Shear, 2016	*The marginal benefits achieved with nicotinamide, however, which appeared sporadically across the TBI models, has reduced enthusiasm for further investigation by the OBTT Consortium.*

Legend: FDP – Freeze-dried plasma; **LP** – Lyophilized plasma; **EPO** – Erythropietin; **OBTT** – Operation Brain Trauma Therapy; **LEV** – Levetiracetam; **FPI** – Fluid percussion brain injury; **CsA** – cyclosporin-A; cyclosporine; **TBI** – Traumatic Brain Injury

Appendix 11: Risk of bias for other sources of bias

Study	Funding influences	Conflicts of interest	Contamination	Unit of analysis errors
Reimer, 1985	L	U	L	U
Crabbe, 1999	L	U	L	H
Alam, 2009	L	U*	L	U
Spoerke, 2009	L	U*	L	U
Jones, 2015	L	L	L	L
Llovera, 2015	L	L	L	U
Maysami, 2015	L	H	L	U
Bramlett, 2016	L	H	L	U
Browning, 2016	L	H	L	U
Dixon, 2016	L	H	L	U
Gill, 2016	L	L	L	U
Mountney, 2016	L	H	L	U
Shear, 2016	L	H	L	U

Source of funding: Low risk = Non-industry source of funding (or no funding). Unclear = Funding source was not reported. High risk = Study was funded by industry.
Conflict of interest: Low risk = Authors reported no conflict of interest. Unclear = Conflict of interest was not reported. High risk = Authors reported on potential conflict of interests.
Contamination: Low risk = No treatment or drug other than the study drug used. Unclear = Possibility of contamination from other treatments or drugs. High risk = Animals receive additional treatment/drugs other than the intervention. Authors report this could influence the results.
Unit of analysis errors: Low risk = Individual units were analyzed individually by the same unit of the treatment comparison group. Unclear: unclear if animals were analyzed individually and treated as one replication. High risk = Units used in the analysis are different from the units of allocation to the treatment comparison groups. Example: animals were not analyzed individual (ex. all animals in one cage) or not treated as one replicate (ex. Same animal: one eye intervention, one eye control).
*financial disclosure, no statement of other conflicts provided

Appendix 12: Completeness of reporting evaluation for 13 studies across 29 items

#	Item Description	Reimer, 1985	Crabbe, 1999	Alam, 2009	Spoerke, 2009	Jones, 2015	Llovera, 2015	Maysami, 2015	Bramlett, 2016	Browning, 2016	Dixon, 2016	Gill, 2016	Mountney, 2016	Shear, 2016
1	Identification as a multicenter study in title*	No	No	Yes	No	Yes	Yes	Yes	No	No	No	Yes	No	No
2	Abstract states number of participating centers	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No
3	Community based reporting guidelines listed	No	No	No	No	No	Yes	Yes	No	No	No	No	No	No
4	Names of each participating center listed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
5	List roles of participating centers (central coordinating center, experimental site)	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
6	No changes, or major changes to study protocol after commencement are documented	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
7	Results substantiated by repetition under a range of conditions at each site	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
8	Number of subjects per outcome	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
9	Number of measurements per subject for one experimental outcome stated	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes
10	Number of subjects per center	Yes	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes
11	List of the total number of subjects used in each experimental group	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes
12	List of all statistical tests used	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
13	Definition of the measure of central tendency	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
14	Definition of the measure of dispersion	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
15	Random group assignment reported	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
16	Description of the method of random group assignment	No	No	No	No	Yes	Yes	Yes	No	No	No	Yes	No	No
17	Experimenters blinded to group allocation during conduct of the experiment	Yes	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
18	Experimenters blinded to group allocation during result assessment	Yes	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
19	Description of an *a priori* primary outcome	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
20	Sample size for each site computed during study design	No	Yes	No	No	Yes	Yes	No	No	No	No	Yes	No	No
21	Description of the method of sample size determination	No	Yes	No	No	Yes	Yes	No	No	No	No	Yes	No	No
22	Total number of animals used for the experiment reported	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	Yes	No	No
23	Description of the criteria used for the exclusion of any data or subjects	Yes	Yes	Yes	No	Yes	Yes	Yes	No	No	No	No	No	No
24	List losses and exclusions of animals at the end of experiment	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No
25	All outcomes described, or description of outcomes measured but not reported in results	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
26	Previous or pilot/preliminary studies performed and listed	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
27	Results were significant, or if not, null or negative outcomes included in the results	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
28	Limitations of the study are documented	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
29	Discrepancies in results across centers expected or absent, or if not, they discussed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	% of reported items per paper	72	83	79	69	97	100	72	69	69	69	66	69	69

***Multicenter or synonym (multi – site, laboratory, institution; cross-site, center, laboratory; consortium). Collaborative or joint study does not meet criteria.**