DEVELOPMENT, VALIDATION AND PILOT IMPLEMENTATION OF THE MINIPIERS (PRE-ECLAMPSIA INTEGRATED ESTIMATE OF RISK) CLINICAL RISK PREDICTION MODEL

by

Beth Payne

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES
(Reproductive and Developmental Sciences)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

December 2014

Abstract

The hypertensive disorders of pregnancy (HDPs) are one of the leading causes of maternal death and morbidity in low-resourced countries due to delays in case identification and a shortage of health workers trained to manage these disorders. The objective of this thesis was to develop an evidence-based tool that could aid community-based health workers in decision making around the care of women with the HDPs.

This objective was achieved using a prospective cohort of data collected in five low and middle income countries (LMICs) to: (1) develop a clinical risk predication model using logistic regression (the "miniPIERS" model); (2) validate the miniPIERS model through bootstrapping and by applying the model to a second cohort of women with HDP; (3) extend and recalibrate the model to include the novel biomarker, pulse oximetry (SpO_2); and (4) translate the miniPIERS model into a decision rule for final creation of the PIERS on the Move decision algorithm. All stages of development of the PIERS on the Move tool included input from stakeholders in low-resourced countries.

The miniPIERS model, based on demographics, symptoms and clinical signs, accurately identified women who were at greatest risk of complications from the HDP (AUC ROC 0.77 [95% CI 0.74 – 0.80]). Internal validation demonstrated minimal overfitting with an average optimism of 0.037. Addition of SpO2 to the miniPIERS model resulted in a 20% increase in classification accuracy of high-risk women. Using an iterative review and feedback process including stakeholders from our partner low-resourced countries, decision points defined by the

miniPIERS model were combined with the WHO recommendations for treatment of women with HDP to create a novel decision algorithm for population level risk screening. This decision algorithm identified high-risk women in the miniPIERS cohort with a sensitivity of 74.1% and specificity of 51.4%. Pilot testing of this tool in South Africa demonstrated potential impact but the true impact of use of the PIERS on the Move tool on maternal outcome rates requires assessment through an implementation study.

Preface

All data on which this dissertation is based were collected as part of the multicentre international miniPIERS (Pre-Eclampsia Integrated Estimate of RiSk) collaborative study. I contributed intellectually to the development of the miniPIERS protocol and data collection instruments and was responsible for primary study coordination. The miniPIERS study was approved by the UBC Clinical Research Board (H07-02207) and the research ethics boards of all partner institutions.

A portion of the introductory chapter and Chapter 2 of this dissertation was adapted from a previous review publication of which I was the lead author and solely responsible for the design and interpretation of the review: Payne B, Magee LA, von Dadelszen P. Assessment, Surveillance and Prognosis in Pre-eclampsia. Best Practice & Research Clinical Obstetrics & Gynaecology (2011) doi:10.1016/j.bpobgyn.2011.02.003 (PMID: 21459048). P von Dadelszen and L Magee edited and approved the final version of the manuscript.

A version of Chapter 3 was published as: Payne B, Hutcheon JA, Ansermino JM, Hall DR, Bhutta ZA, Bhutta SZ, Biryabarema C, Grobman WA, Groen H, Haniff F, et al. (2014) A Risk Prediction Model for the Assessment and Triage of Women with Hypertensive Disorders of Pregnancy in Low-Resourced Settings: The miniPIERS (Pre-eclampsia Integrated Estimate of RiSk) Multi-country Prospective Cohort Study. PLoS Med 11(1): e1001589. doi:10.1371/journal.pmed.1001589 W. (PMID: 24465185). I was responsible for the design of this study including the miniPIERS protocol and data collection instruments, development of

methods used for statistical analysis, performing data analysis, interpretation of results and writing the initial draft of the publication. P von Dadelszen is the principal investigator of the miniPIERS study and was responsible for the initial research concept. J Hutcheon and H Groen contributed to development of study methods and interpretation of results. All co-authors contributed to study design and protocol review as well as edited and approved the final version of the manuscript.

A version of Chapter 4 has been accepted as Payne B, Hutcheon JA, Dunsmuir D, Cloete G, Dumont GA, Hall DR, Lim J, Magee LA, Sikanadar R, Qureshi R, Ansermino MA, von Dadelszen P. (2014) Assessing the incremental value of blood oxygen saturation (SpO2) in the miniPIERS (Pre-eclampsia Integrated Estimate of RiSk) risk prediction model. JOGC (manuscript #14-08-147). I contributed to this study through design of the protocol and data collection instruments and was responsible for all data analysis, interpretation and writing the first draft of the manuscript. All other co-authors also contributed to study design and reviewed and approved the final version of the manuscript. P von Dadelszen and M Ansermino were the study Principal Investigators and were responsible for initiating the project through receipt of funding through the Saving Lives at Birth grant competition with Grand Challenges Canada.

I was the lead investigator, under the direction of my thesis supervisor P von Dadelszen and committee members: J Hutcheon, L Magee and M Ansermino, for all other work presented in this dissertation. As such I contributed intellectually to the design and concept development of the studies presented in Chapters 4 and 6 including writing first drafts of the study

protocols and developing data collection instruments. I completed all data analysis and initial interpretation of results.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AUC ROC	Area Under the Receiver Operating Characteristic curve
ASH	American Society of Hypertension
AST	Aspartate aminotransferase
ALT	Alanine aminotransferase
CHWs	Community-based health workers
DCE	Discreet Choice Exercise
dBP	diastolic Blood Pressure
GA	Gestational Age
GCS	Glasgow Coma Score
GH	Gestational Hypertension
HELLP	Hemolysis Elevated Liver enzymes and Low Platelets
HDP	Hypertensive Disorders of Pregnancy
ISSHP	International Society for the Study of Hypertension in Pregnancy
LDH	Lactate dehydrogynase
LMIC	Low- or Middle- Income Country
LR-/+	Positive or negative likelihood ratio

mHealth	Mobile health
NICE	National Institute for health and Clinical Excellence, UK
NPV	Negative Predictive Value
NRI	Net Reclassification Index
OR	Odds Ratio
PETRA	Pre-eclampsia Trial Amsterdam
PIERS	Pre-eclampsia Integrated Estimate of RiSk
PPV	Positive Predictive Value
POM	PIERS on the Move
RCT	Randomized Controlled Trial
RR	Relative risk Ratio
SOGC	Society of Obstetrics and Gynecologists Canada
SOMANZ	Society of Obstetric Medicine of Australia and New Zealand
sBP	systolic Blood Pressure
Spot Pr/CR	Spot urinary protein to creatinine ratio test
SpO_2	Blood oxygen concentration by pulse oximetry
Sens	Sensitivity

Spec	Specificity
SQI	Signal Quality Index
WHO	World Health Organization

Acknowledgements

There are many people who have contributed to the work presented in this thesis and who are deserving of acknowledgement. First and foremost I would like to thank my supervisor, Dr. Peter von Dadelszen, and the members of my thesis committee, Drs. Jennifer Hutcheon, Mark Ansermino and Laura Magee. Their support and guidance throughout the past many years has been instrumental in my and the project's success. I would also like to acknowledge the support and collaboration of the remaining members of both the miniPIERS and CLIP working groups. I feel extremely lucky and grateful to have been able to work with such an incredible, large and diverse group of researchers, clinicians and health workers.

I gratefully acknowledge the grant support for the PIERS project from the Bill and Melinda Gates Foundation, WHO, Canadian Institutes for Health Research, Pre-eclampsia Foundation, FIGO and Rockefeller Foundation.

Finally, and importantly, I must acknowledge and thank the contribution of the thousands of women around the world that participated in these studies. Without their willingness to contribute to the PIERS cohort, none of this would be possible.

Chapter 1: Introduction

1.1 Classification of the hypertensive disorders of pregnancy

The hypertensive disorders of pregnancy (HDP) include chronic hypertension, gestational hypertension and pre-eclampsia. All disorders are generally defined by the presence of blood pressure greater than or equal to 140/90 mmHg during pregnancy, whether pre-existing (chronic) or of new onset (gestational). Pre-eclampsia is most commonly defined as presence of high blood pressure with proteinuria after 20 weeks gestational age and this is the definition adopted for all studies presented throughout the following chapters.

A number of national and international professional societies have published guidelines on the classification and management of the HDP [1-4]. Unfortunately, consensus amongst these documents relating to definitions and management strategies is lacking. All guidelines include criteria for defining severity of disease. These markers of severity are generally defined by expert opinion and reflect understanding of the effect of endothelial dysfunction on multiple organ systems within the context of the HDP. In addition, many of the criteria for severity stipulated in these guidelines have been shown to have poor prognostic value [5,6]. Table 1-1 provides a summary of the classification systems presented by the UK's National Institute for health and Clinical Excellence (NICE) [2], the Society of Obstetricians and Gynaecologists of Canada (SOGC) [1], the American Society of Hypertension (ASH)[3] and the Society of Obstetric Medicine of Australia and New Zealand (SOMANZ) [4], listing severity criteria and definitions for the HDP.

Definitions of the HDP are subdivided into 1) chronic hypertension; 2) gestational hypertension, and; 3) pre-eclampsia. Pre-eclampsia is then further defined as severe or non-severe, often based on the presence or absence of a variety of additional signs, symptoms and laboratory findings. Unlike the other three guidelines that define severity of disease based on thresholds of biomarker perturbations, in the recently updated SOGC guideline, severity is defined based on occurrence of severe complications that would warrant immediate delivery[1]. This novel way to approach the concept of risk was partially driven by work described in this dissertation.

All of the guidelines reviewed are specific to high-resourced settings and may be difficult to apply to practice in a low- or middle- income country (LMIC) setting because monitoring and treatment processes recommended require significant resource use. For example, severity is often defined based on laboratory parameters that are often not be available to women in LMICs, or are only available at the highest health facility level. The WHO published recommendations for prevention and treatment of pre-eclampsia and eclampsia that are meant to be applicable to low-resourced settings[7]. This document was not included in the guideline review as it does not provide a specific classification system to be used to guide clinical management or define disease severity. The WHO recommendations provide a review of the evidence of all possible treatment options once a diagnosis and severity classification is established.

1.2 Global incidence of the hypertensive disorders of pregnancy

The HDP affect an estimated 5-10% of all pregnancies [8], globally. A recent systematic review of regional estimates of pre-eclampsia and eclampsia, specifically, found that rates of pre-eclampsia ranged from 1.0% (0.1 – 2.6%) in the Eastern Mediterranean region to 5.6% (3.6-11.3%) in the African region while rates of eclampsia ranged from 0.1% (0.0-0.4%) in the Western Pacific and European regions to 2.9% (1.4-7.4%) in the African region. This review included 129 studies from 44 countries reported between 2002 and 2010 of which 52 reported on pre-eclamsia and 42 reported on eclampsia[9]. Studies have suggested rates of HDP are increasing in high-resourced settings such as the United States [8] but due to the lack of accurate historical and population level data in many low-resourced settings, it is unclear if this trend is applicable worldwide.

1.3 Impact on maternal and child health in low-resourced settings

The HDP, and in particular pre-eclampsia and eclampsia, are a significant contributor to the global burden of maternal and perinatal mortality [10-13]. The majority of deaths associated with HDP occur in LMICs in the absence of a trained health professional [14-16]. The increased burden of adverse outcomes in LMICs is believed to be due primarily to delays in triage (identification of who is, or may become, severely ill and should seek a higher level of care), transport (getting women to appropriate care) and treatment (provision of appropriate treatment such as magnesium sulphate, antihypertensives and timed delivery) [17,18]. A major contributing factor to

the morbidity and mortality associated with pre-eclampsia is the shortage of health workers adequately trained in the detection and triage of suspected cases [19].

Maternal morbidities associated with the HDP are thought to be a result of excessive inflammation and endothelial damage [20] and include eclampsia, stroke, retinal detatchment, acute renal failure, placental abruption, pulmonary oedema, liver haematoma, disseminated intravascular coagulation and cerebrovascular bleeding [20-24]. Adverse fetal outcomes include stillbirth, oligohydramnios and fetal growth restriction [20,23]. Both maternal and fetal outcomes tend to cluster around the diagnosis of pre-eclampsia (gestational hypertension and proteinuria) but gestational hypertension alone and other atypical forms of the disorder are not benign [25-29]. Studies have found that 15-56% of women who initially present with gestational hypertension will progress to a diagnosis of pre-eclampsia [25,27,30]. It is estimated that 15% of cases of severe pre-eclampsia, however defined, will result in significant maternal morbidity [31]. The variability in presentation and progression of the HDP presents a significant challenge for the effective management of the disorders.

The most recent report from the Global Burden of Disease study showed an improvement in regards to number of maternal deaths attributed to the HDP between 1990 and 2010 (69,800 vs. 47,100) [10] suggesting an improvement in our ability to care for these women. This trend towards a reduction in total maternal deaths associated with the HDP has also been shown by the WHO [11]. Despite this apparent improvement in mortality the HDP, specifically pre-ecalmpsia, remain one of the top four causes of maternal mortality and morbidity in high,

middle, and low income countries. Using data from 29 LMICs participating in the WHO multicountry survey on maternal and neonatal health, Abalos et al. [32] demonstrated that the odds of maternal death associated with the diagnosis of pre-eclampsia is 3.73 (95% CI 2.15 – 6.47) when compared with women not diagnosed with pre-eclampsia. This risk increased significantly after diagnosis of eclampsia (OR 42.38; 95% CI 25.14 – 71.44). The risks of stillbirth and neonatal death were also significantly increased in women diagnosed with either pre-eclampsia or eclampsia in this study with odds ratios of 3.02 (95% CI 2.73 – 3.34) and 4.91 (95% CI 4.08 – 5.91), respectively [32].

1.4 Treatment and management of the HDP

Despite recent advances in understanding of the pathophysiology of pre-eclampsia [see the recent reviews by Steegers et al [21] and Staff et al [20]], delivery of the placenta remains the only means by which to initiate resolution of the maternal disorder. There are few effective treatments available for managing the disorder after diagnosis has been made. There is strong RCT evidence that magnesium sulfate ($MgSO_4$) is an effective drug for both prevention and treatment of the seizures of eclampsia [33-36]. There are also several reviews that have addressed the safety and efficacy of various antihypertensive medications in pregnancy as a means of controlling blood pressure in hypertensive pregnant women as described in the recently updated SOGC guidelines [1]. Use of antihypertensives in women with HDP has been strongly recommended by the WHO to reduce risk of maternal mortality and morbidity [7]. Neither $MgSO_4$ nor antihypertensive medications will initiate resolution of the HDP. No other pharmacological

or lifestyle intervention has been recommended to treat the HDP once diagnosis has been made [1,2,7].

When presenting early in gestation, delivery is not always the best option for the fetus. Iatrogenic preterm delivery is associated with increased risks of short- and long- term morbidity such as respiratory distress or neurodevelopment delays, whether it occurs in the early or late preterm period [37,38]. Management of the HDP requires balancing the risks to both mother and baby. There is evidence from both cohort studies and RCTs that, remote from term, prolongation of pregnancy by expectant management decreases serious perinatal morbidity without significant increases in maternal risk [39-47]. However, uncertainty remains around the magnitude of the maternal risk associated with expectant management, as initial RCTs were limited to women with severe pre-eclampsia and were underpowered to detect a difference in outcomes between groups [43,44,46,47]. In addition, only one of these studies was performed in a low-resourced setting (reported in two publications) [46,47] so it is unclear how best to apply the results to clinical care in these settings.

More recently research has focused on defining the optimal timing of delivery so that both maternal and fetal outcomes may be optimized in women with non-severe HDP. A recent RCT comparing the impact of routine induction vs. expectant management on maternal outcomes in 756 women (377 in induction arm; 379 in expectant management arm) with gestational hypertension or mild pre-eclampsia recruited between 36-41 weeks gestation concluded that induction of labour at 37 weeks gestation was not associated with increased maternal adverse

events [48]. This RCT was followed up with a second study by the same group to address the impact of routine induction at earlier gestations (34-37 weeks) on both maternal and neonatal outcome as it is during this critical time period that management decisions may have the most significant effects on the long-term health of the baby. In this second RCT inclusion was again limited to cases on mild pre-eclampsia. Results showed that at these earlier gestational ages there was no difference in risk for the mother with routine induction but there was a possible increased risk of neonatal respiratory complications. The authors conclude that the study does not support use of routine induction in the late preterm period for women with mild HDP (unpublished data from personal communication). Observational study data has also been shown to support delivery at 38 or 39 weeks gestation in women with pre-existing hypertension as this was the point at which optimal trade-off between maternal and fetal risks could be demonstrated [49].

Although populations of women from low-resourced settings were well represented in studies of $MgSO_4$, most other research on treatment or management of women with HDP has been performed in high-income facility settings hindering direct application of these strategies to low-resourced settings due to lack of infrastructure and resources. The effect of these resource limitations, as a limitation to effective implementation of expectant management strategies was highlighted in a study by researchers in Egypt where no maternal or perinatal benefit could be drawn from expectant management remote from term due to systematically high morbidity and mortality for babies born below a gestational age of 34 weeks [50]. This reflects the general lack of NICU level care in low-resourced settings that can support babies born preterm,

therefore, delaying delivery for perinatal benefit at early gestation likely only results in increased risk to the mother as the disease progresses. Novel strategies are required that consider this resource gap in order to improve both maternal and fetal outcomes.

1.5 Strategies to improve care of women with HDP in low-resourced settings

As discussed above, management of the HDP depends on the health worker's ability to accurately identify women at greatest risk of developing serious complications (risk stratify) so that interventions can be provided in time and appropriately. As with all causes of maternal mortality, in low-resources settings the consequences of the HDP are made significantly worse by a lack of trained health care workers who can manage this disorder effectively [14,15,51-53]. This health worker shortage results in the "three delays", a theory first proposed by Thaddeus and Maine in 1994 to explain why maternal death is so much more prevalent in LMICs [18]. The first delay focuses on triage and relates to the ability of the care provider and patient to correctly identify the problem and its severity, the second delay is focuses on transport and occurs when there is an inability to access appropriate care either due to lack of transport infrastructure or funds, and the third delay focuses on treatment and describes problems that exist in women receiving appropriate care once the health system is accessed, such as lack of available treatments or staff able to provide proper treatment. These delays interact with the effect of inequities as described by the social determinants of health, making health outcomes worst for the most poor and impoverished on a global level but also within countries [17,54]. Any strategy

aimed at improving maternal health in low-resourced settings must consider how to overcome these delays and the underlying causes of health inequities and health worker shortages.

1.5.1 Task-shifting pregnancy care to community-based health workers

Several strategies have been proposed to reduce the health workforce shortage in LMICs and, therefore, address aspects of the delays in triage, transport and treatment that result in increased maternal mortality and morbidity [14,53,55]. Those that are particularly applicable to the research described in subsequent chapters of this thesis are based on the concept of task-shifting aspects of pregnancy care to community-based health workers [15,19,51,56-58]. Evidence to support the effectiveness of task-shifting is building with several examples of successful use of community-based health workers to provide antenatal and postnatal care to reduce stillbirth and neonatal death rates [15,59] as well as to improve utilization of formal health services [60]. The impact of task-shifting on maternal health outcomes is not as clear but does show promise. Many studies evaluating use of community-based health workers for provision of pregnancy care were underpowered to assess impact on maternal mortality and morbidity alone. A recent systematic review including 10 studies that reported on impact of the intervention on maternal mortality and morbidity found no statistically significant impact on maternal mortality alone (RR 0.77; 95% CI 0.59 – 1.02) but did show a significant reduction in maternal morbidity (RR 0.75; 95% CI 0.61-0.92). This same review also reported on 12 studies including 136,425 pregnancies that showed significant reduction on neonatal mortality (RR 0.76; 95% CI 0.68 – 0.84) and stillbirth (RR 0.84; 95% CI 0.74 – 0.97). For this reason the authors concluded that

there is sufficient evidence to consider scaling up interventions in LMICs that involve task-shifting pregnancy care to community-based health workers [58]. Similar findings are reported in a Cochrane systematic review that demonstrated potential impact on neonatal and child mortality and morbidity but reported insufficient evidence was available to assess impact on maternal morbidity or mortality [60].

Although the evidence base is still building to support the effectiveness of task-shifting for improved maternal and child health, in 2012 the WHO released recommendations for implementation of task-shifting strategies to improve access to essential maternal and child health interventions [56]. This document is meant as a guideline for policy makers and programme implementers and addresses some of the identified barriers to implementation of these programme, such as lack of necessary supervision and stakeholder support [61,62]. Within the guidelines, use of lay health workers to deliver both health education through antenatal visits to women's homes and distribution of essential medicines such as calcium and misoprostol are supported [56]. This was based on findings from a WHO review of country programs utilizing community-based health workers in primary care [63]. This demonstrates the international community's belief in the ability of this cadre of health workers to take on responsibility for some medical care. Criticisms of this strategy have been outlined and include the potential to overburden the existing health workers if too many tasks are assigned and a lack of health resource infrastructure to support and supervise the additional utilization of lower level cadres of health workers should they be given new and more complex tasks [14,19,61,64].

An additional criticism of many of these community-based health workers initiatives is well described by Pakenham-Walsh [55,65-67] in that most of these initiatives have been run based on push paradigms that can be described as research-led, evidence-led and/or subject-led. He argues that most large global health initiatives in LMICs are designed based on research agendas of funding or academic institutions and often fail to take into account the needs of the health workforce and system the initiative is meant to improve. By designing needs-led initiatives that take into account the technological and information needs of the existing health workers, the global health community may have greater and more sustainable impact.

1.5.2 Mobile technology for improved maternal and child health

Use of mobile technology for health (mHealth) has also been suggested as a way to not only support task-shifting but to improve antenatal and postnatal care delivery overall by providing needed technological and information support at the point of care to the health workers who most need it [54,68,69]. The utilization of mHealth in maternal health research in LMICs has quickly developed over the past 5-10 years as a strategy for improved service delivery. This is based on the ubiquity of mobile technology in these settings. At the end of 2013 there were 6.8 billion mobile phone subscriptions active worldwide with 63/100 people in Africa reporting active use of a mobile phone as compared with only 16/100 people having broadband internet access in the same region [70]. This represents an opportunity to the health care system to make use of existing and familiar technology infrastructure to improve care.

Although the promise of impact from mHealth initiatives has been much supported based on anecdotal evidence, few studies have yet to demonstrate impact on real health outcomes. There have been three recent systematic reviews assessing the impact of mHealth programmes on some aspect of maternal or neonatal care that have included studies performed in low-resourced settings [69,71,72]. In all cases the authors highlighted the limited high-quality evidence of effectiveness. High-quality in this instance is relating to methodology of the study performed (for example use of randomization) and outcome assessed. The review by Braun et al [71] focuses specifically on the use of mobile technology to support community-based health workers. The review included 25 published articles that assessed a range of outcomes including quality of care indicators, utilization of health services and quality of health data collection. The authors conclude that preliminary results from the programmes evaluated and included in the report demonstrate the potential for mobile technology to support improved health services through community-based health workers but again highlights the need to properly evaluate these programmes in light of impact on health outcomes [71]. This type of impact evidence is required to ensure cost-effectiveness and sustainability of programmes from a policy-makers perspective [68].

1.6 Summary

The HDP are common and present a significant burden to health systems, women and their families in low-resourced settings. Many guidelines for classification and management of these disorders exist. Many risk factors and severity criteria included in these guidelines are based on

expert opinion and have failed to show adequate performance in prognostic studies, as described in the next chapter. In addition, many of the risk factors included in these guidelines require high cost and laboratory facilities to complete their assessment, making them unsuitable in a low-resourced setting.

Several strategies for improving care of pregnant women in low-resourced settings have been proposed. These include task-shifting care to available yet minimally trained community-based health workers and use of mobile technology to assist in this process. Before this can be done, tools that specifically address the gap in community-based health worker knowledge and ability to provide care to women with HDP are required, and evidence to support impact of these tools on health outcomes is needed.

1.7 Thesis objectives

The overall objective of the research presented in this thesis was to develop an evidence-based clinical decision support tool for use in low-resourced settings by minimally trained health workers to aid in management of women diagnosed with a hypertensive disorder of pregnancy (HDP). This research tests the hypothesis that simple demographics, symptoms and signs alone can be used to stratify women with a HDP into higher- and lower-risk groups, which will improve clinical care by reducing the delays in triage, treatment and transport. Ultimately, it is hoped that this tool will reduce the incidence of adverse outcomes associated with the HDP.

This objective is met through several stages of model development and assessment in the subsequent chapters of this dissertation. In Chapter 2 a literature review is presented on

prognostic and risk factors for adverse maternal outcomes in women with HDP. This review was completed to inform the choice of candidate predictors used in the development and validation of the miniPIERS risk prediction model, presented in Chapter 3. In Chapter 4, conversion of the miniPIERS model into a decision rule is described along with the development of a broader decision algorithm on which to base development of a mobile health decision aid for community-based health workers. In Chapter 5, an analysis of the impact of pulse oximetry as a novel biomarker in the miniPIERS model is provided. Finally, Chapter 6 presents results of research into the effect of various stages of the model development process itself on optimism of the final prediction model generated.

Table 1-1: Definitions of HDP and disease severity according to international guidelines

	NICE (2010)	SOMANZ (2008)	SOGC (2014)	ASH (2008)
Pre-existing / chronic hypertension (BP ≥140/90 prior to or before 20^{+0} week's gestation)	Chronic Hypertension - before 20 weeks' gestation or being treated at time of referral - primary or secondary aetiology	Chronic hypertension - essential - secondary - white coat - with/without superimposed pre-eclampsia	Pre-existing hypertension - with/without co-morbid conditions - with/without signs of pre-eclampsia	Chronic hypertension with or without super-imposed preeclampsia

	NICE (2010)	SOMANZ (2008)	SOGC (2014)	ASH (2008)
Gestational hypertension (GH) **(BP ≥140/90 after 19^{+6} weeks' gestation)**	Gestational hypertension without significant proteinuria	Gestational hypertension without significant proteinuria returning to normal within 12 weeks postpartum	Gestational Hypertension - with/without co-morbid conditions - with/without signs of pre-eclampsia	Gestational hypertension: Or transient hypertension; - blood pressure returning to normal within 6 weeks' postpartum; - late postpartum hypertension, with blood pressure rise developing weeks' to 6 months post partum and normalised by 1 year post partum

	NICE (2010)	SOMANZ (2008)	SOGC (2014)	ASH (2008)
Pre-eclampsia (clinical definition)	New hypertension (BP ≥140/90) presenting after 20 weeks' gestation with clinically relevant proteinuria (see significant proteinuria, below)	GH + one or more of the following: Proteinuria: Pr:Cr ratio >30mg/mmol or 0.3g/24hrs; serum or plasma creatinine >90μM; oliguria - thrombocytopenia; haemolysis; disseminated intravascular coagulation - raised serum; transaminases; severe epigastric or right upper quadrant pain - eclampsia; stroke - hypereflexia with sustained clonus; severe headache; persistent visual disturbances - pulmonary oedema - fetal growth restriction; placental abruption	-Pre-existing hypertension or resistant hypertension with new proteinuria, or adverse condition -GH + proteinuria (spot Pr:Cr ratio >30mg/mmol or 0.3g/24h), or adverse condition Adverse conditions: headache/ visual disturbances; chest pain/ dyspnoea; oxygen saturation <97%; aPTT, serum creatinine, serum uric acid, AST, ALT, LDH or bilirubin; low platelet count, plasma albumin; nausea or vomiting, right upper quadrant pain or epigastric pain; non-reassuring FHR; IUGR; Oligohydramnios; absent or reversed end-diastolic flow by Dopper velocimetry	Gestational hypertension or chronic hypertension with proteinuria (dipstick ≥+1, spot Pr:Cr ratio ≥30 mg/mmol or ≥0.3 g/24hrs)

	NICE (2010)	SOMANZ (2008)	SOGC (2014)	ASH (2008)
Pre-eclampsia (research definition)	Not defined	De novo hypertension >20^{+0} weeks', returning to normal postpartum with properly documented proteinuria	Not defined	Not defined
Severe hypertension	160/110 mmHg	170/110 mmHg	160/110 mmHg	160/110 mmHg
Significant proteinuria	>300 mg/d or >30 mg/mmol on spot Pr:Cr ratio	Not defined	>300 mg/d or >30 mg/mmol on spot Pr:Cr ratio	>300 mg/d or >30 mg/mmol on spot Pr:Cr ratio

	NICE (2010)	SOMANZ (2008)	SOGC (2014)	ASH (2008)
Severity criteria	-severe hypertension - maternal symptoms (vision problems, severe headache, epigastric pain, vomiting, papiloedema) - biochemical abnormalities or haematological impairment (platelet count <100 x 10^9/l or AST/ALT >70 U/L, elevated serum creatinine)	Not defined	Occurrence of any of the following severe complications: - eclampsia; PRES; cortical blindness or retinal detachment; GCS<13; Stroke, TIA or RIND; uncontrolled severe hypertension; oxygen saturation <90%; need for >50% oxygen for > 1 hour; intubation other than for caesarean section; pulmonary oedema; myocardial infarction or ischemia; platelet count <50x10^9/L; transfusion of any blood product; acute kidney injury; new indication for dialysis; hepatic dysfunction, rupture or haematoma; placental abruption; reverse ductus venosus A wave; stillbirth	<35 weeks' gestation -maternal symptoms (headache, visual disturbances, abdominal pain) - severe diastolic hypertension (>110 mmHg) - significant proteinuria or oliguria - increased serum creatinine - decreased glomerular filtration rate - increased AST or LDH - fetal morbidity (non-reassuring cardiotogograph)

ALT: alanine transaminase; AST: aspartate transaminase; BP: blood pressure; GA: gestational age; GH: gestational hypertension; HELLP: haemolysis, elevated liver enzymes, and low platelets; LDH: lactate dehydrogenase; PIH: pregnancy-induced hypertension; Pr/Cr: protein-to-creatinine ratio; NICE: National Institute for health and Clinical Excellence; SOMANZ: Society of Obstetric Medicine of Australia and New Zealand; SOGC: Society of Obstericians and Gynaecologists of Canada

Chapter 2: Prognosis in women with hypertensive disorders of pregnancy

2.1 Background

Any disease classification systems, such as the guidelines presented in the previous chapter, are only useful if they can accurately stratify the population of interest based on differences in prognosis and can identify those patients for whom interventions or treatments are most appropriate. Prognosis in medicine refers to the probability of developing an adverse health outcome during the course of that person's care. It is important to distinguish the identification of a risk factor and a prognostic factor. These concepts are often used interchangeably in the medical literature but are actually very different. Risk factors are provided for individual patients as odds ratios or relative risks and describe the causal association between the adverse outcome and the presence or absence of a biomarker or clinical measure. Prognostic factors on the other hand, give information on the absolute probability of an outcome given the presence or absence of the predictor, independent of causation [73-75].

Having the ability to predict the likelihood of an individual developing a poor outcome based on that individual's clinical picture is critical for decision making by both the health care provider and patient [73].

2.2 Prediction of adverse maternal outcomes by individual investigations

A number of studies have assessed the prognostic value of the clinical and laboratory investigations used to define severity in women with pre-eclampsia in at least one of the

guidelines reviewed in section 1.1 of the previous chapter. In this section the evidence to support inclusion of individual measures as severity criteria is reviewed.

2.2.1 Blood pressure

Severe hypertension defined as a systolic blood pressure (sBP) ≥160 and/or diastolic blood pressure (dBP) ≥ 110 mmHg is given in all guidelines reviewed as a severity criterion for pre-eclampsia. The majority of studies found only investigated the association between high blood pressure and adverse outcomes and did not report prognostic value. In one case study including 28 women diagnosed with pre-eclampsia, systolic blood pressure >160 mmHg was shown to be an independent risk factor for stroke [76]. Conversely, in another study including 216 patients diagnosed with HELLP (haemolysis, elevated liver enzymes and low platelets) syndrome or severe pre-eclampsia using the ASH criteria [3], diastolic blood pressure at inclusion >105 mmHg was shown to be associated with a reduced risk of adverse maternal outcomes (OR 0.66; 95% CI 0.45 – 0.96) [40]. In other studies, no association between blood pressure and adverse outcomes was found [5,77]. We also reported the predictive value of blood pressure was in the fullPIERS (Pre-eclampsia Integrated Estimate of RiSk) cohort study of 1259 women diagnosed with pre-eclampsia, as defined by the SOGC [1], and found that neither diastolic blood pressure nor systolic blood pressure adequately predicted risk of a combined adverse maternal outcome alone (AUC ROC dBP: 0.66; 95% CI 0.58, 0.748, sBP: 0.69; 95% CI 0.61, 0.78) [78]. No blood pressure cut-off that predicts risk can be identified based on these results. The poor prognostic value of blood pressure may be due to the fact that this is a highly modifiable variable within

this population through use of effective antihypertensive agents, therefore, masking the true effect of the rise in blood pressure on incidence of adverse outcomes.

2.2.2 Proteinuria

New onset of proteinuria in pregnancies complicated by hypertension has formed the basis of the clinical diagnosis of pre-eclampsia for many years and is included in all international guidelines. Significant proteinuria, defined using various thresholds, has also been used as criteria for severity of disease in the ASH [3] guideline and is listed as an adverse condition requiring greater monitoring in the SOGC guideline [1]. More recently, the role of proteinuria measurement in the classification of pre-eclampsia has been called into question [see the review by Lindheimer and Kanter [79]]. Inaccuracies have been identified with the gold-standard 24-hour urine collection method [80] prompting investigation into the utility of other methods of measuring proteinuria in hypertensive pregnancies. The spot protein:creatinine ratio test is an alternative method of testing proteinuria in this population. Results from a systematic review have suggested that using 30 mg/mmol as a threshold for the spot protein:creatinine ratio is a reasonable alternative to 0.3 g/day to rule-out proteinuria in hypertensive pregnancies (Sensitivity 83.6%; 95% CI 77.5 – 89.7% and specificity 77.5%; 95% CI 72.6-80.0%) [81] but these thresholds were arbitrarily chosen and require further validation.

The occurrence of proteinuria in women with pre-eclampsia has been shown to be associated with some increased perinatal risks but not necessarily maternal risks. In one study including 1348 women diagnosed with pre-eclampsia using the International Society for the Study of

Hypertension in Pregnancy (ISSHP) classification system [82], the occurrence of proteinuric pre-eclampsia was associated with significantly increased odds of preterm birth (OR 1.46; 95% CI 1.11 – 1.92) and perinatal mortality (OR 4.28; 95% CI 1.01 – 18.16) and showed a trend towards more severe hypertension (OR 1.28; 95% CI 0.98 – 1.68) when compared with women with non-proteinuric pre-eclampsia [26]. Using a subset of 321 women included in this study, the prognostic value of proteinuria measured by spot Pr/Cr was investigated. The final model included both spot Pr/Cr ratio and maternal age with an AUC ROC of 0.67 for prediction of maternal adverse outcomes. Maternal age was included because investigators found that increased maternal age reduced the level of proteinuria required to be predictive of adverse outcomes. This same study found that no level of proteinuria could be defined to accurately predict outcomes, leading to the conclusion that the presence of proteinuria rather than magnitude is more reflective of risk [83]. This conclusion was supported by an analysis I performed using 2023 women with pre-eclampsia in the fullPIERS cohort, where none of dipstick, spot Pr/Cr, nor the 24hr urine test accurately predicted maternal or fetal outcomes (AUC ROC <0.7 in all cases) despite a strong association found between significant proteinuria and risk of our combined neonatal outcome (including stillbirth, neonatal death or admission to NICU for greater than 48 hours) [84]. Finally, a systematic review of proteinuria, measured by 24 hour urinalysis, as a predictor of complications in women with severe pre-eclampsia, variably defined, found that proteinuria is a poor predictor of both maternal and fetal complications. Due to the heterogeneity of cut-off values used to define proteinuria, pooling of data was not possible. The adverse maternal outcomes investigated were eclampsia and placental abruption

and resulted in likelihood ratios for a positive and negative test that were below the level required for usefulness in prognosis. For studies included, the positive likelihood ratio (LR+) ranged from 2.7 – 1.7 and negative likelihood ration (LR-) ranged from 0.41 – 0.62 for the prediction of eclampsia. For the prediction of placental abruption, pooled LR+ was 0.88 (95% CI 0.42 – 1.86) and LR- 1.1 (95% CI 0.75 – 1.6). No significant results were found for the prediction of fetal outcomes including stillbirth, neonatal or infant death or NICU admission [85].

2.2.3 Laboratory tests

Several laboratory tests are recommended for the surveillance of women with hypertension in pregnancy. Of these tests, platelet counts <100 x 10^9/L; elevated serum creatinine; elevated aspartate transaminase (AST) or lactate dehydrogenase (LDH); and plasma albumin <20g/L have been used to define severity and are explicitly recommended as indications for delivery by the SOMANZ [4]. This is based on increased incidence of maternal adverse outcomes shown to occur in women with these various criteria, but few studies have addressed whether these tests can be used to predict the probability of adverse maternal or fetal outcomes. Thrombophilic disorders have been shown to be associated with increased risk of the placental disorders such as IUGR and the HDP and it has been suggested that this association is due to effects on platelet function [86-88]. Platelet count has been well studied as a risk factor for severe disease. Several studies including women diagnosed with pre-eclampsia have demonstrated that platelet count <100 x 10^9/L is associated with increased incidence of adverse maternal and perinatal outcomes ($p<0.05$ in all cases) [5,88,89]. In one study, on which I am a co-author, including 1387 women in

the fullPIERS database diagnosed with pre-eclampsia using the SOGC guidelines [1], a platelet count <100 x10^9/L was shown to have poor utility in predicting risk of adverse maternal outcomes (LR+ 4.05; 95% CI 2.60 – 6.31; LR- 0.78; 95% CI 0.67 – 0.90) [90]. No studies have evaluated the prediction of adverse fetal outcomes using platelet count.

Similarly, serum creatinine >110 µM was shown to be associated with increased incidence of adverse maternal outcomes (p<0.001) [5]. Again, using data from the fullPIERS study, including 1259 women with pre-eclampsia, we demonstrated that creatinine was associated with increased risk of adverse maternal outcomes in a univariate logistic regression analysis (OR 1.02; 95% CI 1.02-1.03) but based on the reported area under the curve of the receiver operator characteristic (AUC ROC 0.67; 95% CI 0.57-0.76), creatinine does not perform adequately to be used alone to predict risk [78].

Elevated liver enzymes have been found to be independent risk factors for adverse maternal outcomes (p<0.001) in women with pre-eclampsia [5,91] but not in women with HELLP syndrome [29,92]. The prognostic value of liver enzymes: AST; ALT; and LDH, were investigated by our group using data from 1938 women in the fullPIERS database. The AUC ROC for AST, ALT and LDH were all below 0.7 and no threshold value that predicts risk could be identified [93]. In this study, plasma albumin levels were also investigated and shown to be associated with increased risk of adverse maternal outcomes (OR 2.5; 95% CI 1.4–4.6) when comparing the women in the lowest quartile (albumin <21 g/L) to those in the highest (albumin >41 g/L), but this did not result in an adequately discriminative test [93].

More recently, serum uric acid has been suggested to have clinical utility as a prognostic indicator of maternal health in pre-eclampsia due to its action on the endothelium [94-96]. Serum uric acid levels are known to be elevated in women with pre-eclampsia and gestational hypertension when compared with normotensive pregnant women [97]. One meta-analysis has been published addressing the accuracy of serum uric acid as a predictor or maternal and fetal complication in women with severe pre-eclampsia, variably defined. In a combined cohort of 634 women, using a threshold level of 350 μmol/L, serum uric acid was a poor predictor of eclampsia (pooled LR+ 2.1; 95% CI 1.4-3.5; LR- 0.38; 95% CI 0.18 – 0.81). No other adverse maternal outcomes were investigated [98]. Using the fullPIERS data, our analysis supported the potential utility of uric acid as a prognostic marker after standardization of the values across gestation using a z-score was applied (unpublished). In another study, including 258 women diagnosed with severe pre-eclampsia using criteria similar to those endorsed by the ASH [3], investigators found a moderately increased likelihood of a combined adverse maternal outcome with a positive test using a threshold value of 475.8 μmol/L (LR+ 3.50; 95% CI 1.27 - 9.64; LR- 0.85; 95% CI 0.71 – 1.03) [99]. Due to the heterogeneity of definitions and thresholds used in studies, the clinical utility of uric acid measurement remains unclear.

2.2.4 Pulse oximetry

The predictive value of blood oxygen saturation measured by pulse oximetry (SpO_2) in pregnancies complicated by pre-eclampsia was investigated in one study, on which I am a co-author. In this study, using a prospective cohort of 1534 women in the fullPIERS database,

oxygen saturation was found to be an accurate predictor of risk of a combined adverse maternal outcome within 48 hours of admission to hospital with pre-eclampsia (AUC ROC 0.71; 95% CI 0.65-0.77). Threshold levels for low, medium and high risk were identified as 96-97%, 94-95% and 90-93%, respectively [100].

2.2.5 Symptoms

Clinical symptoms of headache, nausea and vomiting, right upper quadrant or epigastric pain, chest pain or dyspnoea and visual disturbances have all been found to be associated with the HDP and are used to define severity of disease by all but the most recent SOGC [1] guideline. The use of clinical symptoms as prognostic factors in pre-eclampsia is controversial, as many of these symptoms are non-specific and common to normal pregnancy. In a small cohort of 61 women diagnosed with HELLP syndrome, the presence of headache (OR 3.6; 95% CI 1.2 – 10.4), visual disturbances (OR 5.2; 95% CI 1.7 – 15.9), and epigastric pain (OR 3.75; 95% CI 1.04 – 13.4) have all been shown to be associated with increased incidence of maternal adverse outcomes [92]. In a retrospective cohort of 970 women with severe pre-eclampsia, with or without HELLP syndrome, one study found that nausea and vomiting and epigastric pain where associated with increased incidence of adverse maternal outcomes ($p<0.01$), particularly when combined with abnormal laboratory test results for platelet count, liver enzymes (AST, ALT or LDH), serum uric acid and serum creatinine [91]. When I analyzed data from 1259 women included in the fullPIERS database, my co-authors and I showed that only the symptom complex - chest pain and dyspnoea, was predictive of adverse maternal outcomes, although poorly (AUC ROC 0.642; 95%

CI 0.54-0.74) [101]. These data suggest that symptoms are of limited utility for determining risk and should not be used alone to guide clinical decisions.

2.3 Prognosis based on using multivariate clinical risk prediction models

Given the multi-system nature of pre-eclampsia and the variability in presentation, it is not surprising that no individual variable can be identified to predict adverse outcomes alone. Multivariate prognostic models have been developed and successfully implemented in several other areas of medicine (for example: [102-106]). These models, when properly developed and validated, can be used to identify those patients for whom intervention would be most beneficial and can aid in decision making for both the patient and health care provider [73].

2.3.1 Methods for development and validation of multivariate models

Methods for development and validation of clinical risk prediction models have been presented, although several methodological issues remain outstanding. Development of prediction models generally includes variable selection based on either previous knowledge of risk factors, such as that presented in the previous section of this chapter, or by using an automated selection process. Parameter estimation using a regression technique suitable to the outcome of interest is then performed to fit the most optimal model in the development dataset. Data used can be derived from cohort studies or RCTs. Model validation involves both internal and external processes.

Internal validation is performed to determine likely degree of overfitting of the model to the development dataset and optimism in the estimates of model performance measures. Internal

validation techniques allow estimates to be made of the likely decrease in performance that will be seen when the model is applied to an external dataset. Several methods for estimating overfitting and optimism in performance exist. These include use of a split-sample for development and internal validation, cross-validation and the jackknife method or the bootstrap [107,108]. The use of a split-sample involves arbitrary splitting, either in time or through random selection, of the available data to create two independent datasets. Model development is carried out on one portion of the data and then the reserved data is used to estimate model performance measures such as discrimination and calibration [107,109,110]. Although this method is routinely used for model development and internal validation due to its ease of use and understandability, it is widely criticized in the epidemiological literature due to the inefficiency inherent in use of a small proportion of available data for model development and the apparent underestimation of model optimism [111-113].

Cross-validation is similar to data splitting in that model development involves use of a random selection of data from the dataset and then performance is assessed on the remaining data. Unlike data splitting, cross-validation makes use of a majority proportion of the data for development (for example 90%) then uses the small remainder of data for testing model performance. This process is then repeated multiple times so that each case within the original dataset is used for model testing at least once and the performance estimated as the average of all individual cross-validation estimates [107]. The jackknife is the most extreme example of this procedure where one case is left out each time the model is developed and performance assessed based on the average estimated from each N-1 version of the dataset. This method

has been criticized as a means of assessing model validity and the degree of overfitting due to the fact that the development dataset is essentially the same as the full dataset, so assessment of model uncertainty due to automated variable selection methods is unreliable [111,113].

The bootstrap is a resampling technique first developed for use as a nonparametric method for estimating the variability in a parameter of interest or distribution of a test statistic [114]. The bootstrap method differs from that of data splitting and cross validation in that samples are drawn from the original study population with replacement to generate several new study samples of the same size as the original. No data are wasted, making the bootstrap method the most efficient for estimating model optimism [107,115]. The bootstrap method is widely recommended for assessment of internal validation of prediction models due to its efficiency, its ability to account for variable coding and selection processes in the assessment of model performance and the stability of estimates that can be obtained. Previous simulation studies have shown stability of estimates is achieved at 200-500 iterations and no benefit is achieved when using a higher number of repetitions [108,116].

The final steps for validation and testing of any clinical risk prediction model are external validation and implementation [108,117]. External validation requires application of the developed model to a new dataset and population. Generally, a first step is to confirm model performance in a population similar to that used for model development before testing on more distinct populations to assess generalizability of results in broader groups of patients [111,118]. Finally, if

the model performance is maintained, implementation of the model in clinical practice is warranted but should include a process of evaluation on outcome incidence [113,118].

2.3.2 Models developed to determine prognosis in women with a HDP

Two studies have attempted to develop multivariate prognostic models in women with pre-eclampsia. In one study, using a prospective cohort of 216 women admitted as part of the PETRA (Pre-Eclampsia Trial Amsterdam) study with HELLP, severe pre-eclampsia (defined by the ISSHP [82](merged)), eclampsia or fetal growth restriction with gestational hypertension, variables were evaluated based on their ability to predict adverse maternal outcomes at any time after eligibility. Variables found to be associated with the adverse outcome were: estimated fetal weight below 1100g (RR 1.49; 95% CI 1.02-2.18); diastolic blood pressure >105 mmHg (RR 0.66; 95% CI 0.45-0.96); thrombophillic disorders (RR 1.51 95% CI 1.05-2.18); maternal age above 30 years (RR 0.62; 95% CI 0.42-0.92), and; nulliparity (RR 2.19; 95% CI 1.27-3.78). When these variables were included in a step-wise backward elimination logistic regression model building process, the resultant prediction model included only estimated fetal weight and nulliparity and had poor discriminative power (AUC ROC 0.65; no confidence interval reported) [40]. No further analysis on the application of this model in clinical care was justified.

The fullPIERS study, on which I worked as the research coordinator, is a prospective, multicentre observational study that was designed specifically to develop and validate outcome prediction model for women admitted to hospital with pre-eclampsia [119]. Unlike previous studies, inclusion criteria were not limited to women with HELLP or severe pre-eclampsia and

included all women admitted with hypertension and proteinuria, hypertension and hyperuricaemia, HELLP, or superimposed pre-eclampsia (as defined by the SOGC [1,120](merged)). Using a cohort of 2023 women admitted to tertiary academic centres in the United Kingdom, Canada, New Zealand and Australia, variables including demographics, symptoms, signs and laboratory findings were evaluated based on their ability to predict adverse maternal outcomes within 48hours of eligibility or up to 7 days after eligibility [119]. This time point for outcome prediction differs from that used by Ganzevoort et al. [40] and was chosen to allow time for corticosteroid administration and transport.

The final fullPIERS model included: gestational age at onset of disease or delivery (if onset is postpartum) (OR 0.91; 95% CI 0.88 – 0.95) ; serum creatinine (OR 1.02; 95% CI 1.02-1.02); platelet count (OR 0.99; 95% CI 0.98-0.99); AST (OR 1.01; 95% CI 1.00-1.01); SpO_2 (OR 0.63; 95% CI 0.58-0.70), and; chest pain or dyspnoea (OR 6.13; 95% CI 3.56-10.54). This model accurately predicted adverse maternal outcomes within 48hrs of eligibility (AUC ROC 0.88; 95% CI 0.84-0.92) and up to 7 days (AUC ROC 0.76; 95% CI 0.72-0.80) [119].

The original fullPIERS model received criticism because it was developed using predictor variables collected within 48hrs to predict an outcome within the same timeframe [121]. To address this concern, our team subsequently assessed the model in the original dataset using information collected on admission to hospital to confirm predictive performance was maintained in this more clinically relevant timeframe. In all cases the discrimination of the model was only mildly reduced [using predictor variables available within 6 hours of admission

(AUC ROC 0.76, 95% CI 0.72-0.81), and within 24 hours of admission (AUC ROC 0.81, 95% CI 0.77-0.86)], showing that the model could be applied at the time of admission with similar result [122]. Although the fullPIERS model shows promise as a tool to improve health care workers ability to manage women with HDP, given the inclusion of laboratory tests, it is not appropriate for use in a low-resourced setting.

2.4 Summary

The ability to determine accurate prognosis in women with hypertensive disorders of pregnancy could significantly improve both care providers and patient's decision-making ability around the use of interventions versus expectant management. This could not only improve outcomes for the women but has implications on overall health resource use. Although many potential prognostic factors have been identified through their demonstrated association with occurrence of adverse outcomes, none of these factors when assessed alone perform adequately to effect decisions around care.

It is only through the combination of demographics, symptoms and signs that accurate prognosis is possible, as demonstrated by the fullPIERS study[123]. Methods for development and validation of multivariate prognostic models have been defined. The fullPIERS model has been internally validated and appears to perform well but to have the greatest impact in a low-resourced setting, a new model based on available measures and developed using a population of women from the setting in which it is meant to be used is required.

Chapter 3: Development and validation of the miniPIERS (Pre-eclampsia Integrated Estimate of RiSk) model for mothers in low-resourced settings

3.1 Introduction

As described in the previous chapter, our group has previously developed the fullPIERS (Pre-eclampsia Integrated Estimate of RiSk) clinical prediction model, which predicts adverse maternal outcomes among women with pre-eclampsia based on a woman's gestational age at diagnosis, the symptom complex of chest pain and/or dyspnoea, oxygen saturation by pulse oximetry, and laboratory results of platelet count, serum creatinine, and aspartate transaminase. The fullPIERS model, validated in a high income tertiary hospital setting, has excellent discriminatory ability with an area under the receiver operating characteristic curve (AUC ROC) of 0·88 (95% CI 0·84 – 0·92) [119]. However, due to the inclusion of laboratory tests, the fullPIERS model may not be suitable for all settings, particularly primary care settings in LMICs.

The objective of the miniPIERS study described in this Chapter was to develop and validate a simplified clinical prediction model for adverse maternal outcomes among women with HDP for use in community and primary health care facilities in LMICs. This model was intended to be used as a decision aid in the field, allowing community based health workers to more effectively identify and manage cases of pre-eclampsia. By identifying those women at highest risk of adverse maternal outcomes well before that outcome occurs, transportation and treatment can be targeted to those women most in need.

3.2 Methods

3.2.1 Study design and population

The miniPIERS model was developed and validated on a prospective, multicentre cohort of women admitted to a participating centre with a HDP. Participating institutions were: the Colonial War Memorial Hospital, Suva, Fiji; Mulago Hospital, Kampala, Uganda; Tygerberg Hospital, Cape Town, South Africa; Maternidade Escola de Vila Nova Cachoeirinha, São Paulo, Brazil; Aga Khan University Hospital and its secondary level hospitals at Garden, Karimabad and Kharadar; and the Jinnah Post-graduate Medical College, Karachi, Pakistan; and Aga Khan Maternity & Child Care Centre, and Liaqat University of Medical Sciences, Hyderabad, Pakistan. Ethics approval for this study was obtained from each participating institutions research ethics board as well as the clinical research ethics board at the University of British Columbia. All participating institutions had a hospital policy of expectant management for women with pre-eclampsia remote from term, and similar guidelines for treatment of women with regard to magnesium sulphate and antihypertensive agents. Institutions were chosen to participate based on the consistency of these guidelines in order to achieve some level of homogeneity within the cohort and to reduce systematic bias that could result from differences in disease modifying practices between institutions.

Women were admitted to the study with any HDP defined as follows: pre-eclampsia, defined as i) blood pressure (BP) ≥140/90mmHg (at least one component, twice, ≥4 and up to 24 hours apart, after 20 weeks) and either proteinuria (of ≥2+ by dipstick, ≥300mg/d by 24 hour

collection, or ≥30 g/mol by urinary protein:creatinine ratio) or hyperuricaemia (greater than local upper limit of local non-pregnancy normal range), ii) HELLP (haemolysis, elevated liver enzymes, and low platelets) syndrome even in the absence of hypertension or proteinuria [124], or iii) superimposed pre-eclampsia (clinician defined rapid increase in requirement for antihypertensives, systolic BP (sBP) ≥170mmHg or diastolic BP (dBP) ≥120mmHg, new proteinuria, or new hyperuricaemia in a woman with chronic hypertension); or an 'other' HDP defined as: i) gestational hypertension ((BP) ≥140/90mmHg (at least one component, twice, ≥4 hours apart, ≥20^{+0} weeks) without significant proteinuria); ii) chronic hypertension (BP ≥140/90mmHg before 20^{+0} weeks' gestation); or iii) partial HELLP (i.e. haemolysis and low platelets OR low platelets and elevated liver enzymes). All women participating in the study gave informed consent according to local ethics board requirements.

Women were excluded from the study if they were admitted in spontaneous labour, experienced any component of the adverse maternal outcome before eligibility or collection of predictor variables, or had confirmed positive HIV/AIDS status with CD4 count <250 cells/mL or AIDS-defining illness.

3.2.2 Candidate predictors

Candidate predictor variables for final model development were identified *a priori* as being those variables that were: (a) available and easy to collect in all health care settings including the woman's home; (b) associated with pre-eclampsia in previous studies [6](merged); and (c) measurable using simple and reliable methods. These variables included demographics

(maternal age, parity, and gestational age on admission); symptoms (headache, visual disturbances, chest pain/ dyspnoea, right upper quadrant pain or epigastric pain, nausea, vomiting, and vaginal bleeding with abdominal pain); and signs (blood pressure, and dipstick proteinuria). A copy of the data collection form is attached as Appendix A at the end of this dissertation. The values for these variables were collected prospectively from the woman's medical record as measured by the nurse or physician during regular antenatal, intrapartum or postnatal care. If multiple measures of a candidate predictor were collected within the first 24 hours of admission, the worst predictor value obtained within that first 24hrs of admission was used. The value used was the worst in the clinical context, this could either be the highest or lowest value collected in the given 24hr time period, depending on the measure in question. This method of using the worst value was chosen as it is consistent with clinical practice. Generally, clinicians will respond to the worst clinical value when making management decisions.

The external validation study was performed using data from the fullPIERS [119] dataset. Participating centers were tertiary academic hospitals located in Canada (6), the UK (2), New Zealand (1) and Australia (1). Only the fullPIERS data collected after March 1, 2008 were used for this study as this portion of the fullPIERS cohort was collected using the same protocol, inclusion and exclusion criteria and data collection tools as later used for miniPIERS. Prior to this date, the fullPIERS cohort did not include abdominal pain, vaginal bleeding or any headache.

3.2.3 Main outcome measures

The components of the composite adverse maternal outcome to be predicted by the model were determined by Delphi consensus [125] and include maternal mortality or one or more of serious central nervous system, cardiorespiratory, renal, hepatic, haematological or other morbidity. A full list of outcome components and definitions is provided as Appendix B. The Delphi consensus process involved iterative review and feedback on the proposed outcome components from an expert group consisting of researchers and clinicians from both high and low- or middle- income countries who have published work focused on the HDP, giving them clinical and content expertise. Representatives of the Delphi group brought expertise from Medicine, Obstetrics, Paediatrics, Anaesthesia and Critical Care with sub-specialty expertise in Maternal-Fetal Medicine, Nephrology, Haematology, and Placental biology. Data were collected on the occurrence of all outcome components at any time during admission but for the purpose of the model, only those that occurred within 48 hours of admission were considered. All study sites were instructed to collect information on any "other" adverse events the woman experienced during pregnancy or immediately postpartum as part of the regular data collection process. This was done to ensure balanced reporting of events across all sites. Any reported "other" events were adjudicated by the study Working Group during regular meetings, at which time the decision was made whether to include the reported outcome as a study outcome, or not.

3.2.4 Data quality and missing data

Data for the miniPIERS dataset were collected prospectively using standardized data collection forms and protocols for all sites and entered into a customized Microsoft Access database. As part of the study protocol, women were required to have at least one measure of proteinuria, blood pressure and symptoms during the first 24 hours of admission. All data were reviewed for quality and consistency. When questions arose regarding data, these data were confirmed by re-review of the primary health record. Random review of 10% of cases was performed during the first year of the study to ensure data validity within and between study sites.

3.2.5 Sample size

The sample size required for model development was determined based on the minimum standard of 10 events per effective variable considered in the model according to the formula N=(nx10)/I where N is the sample size, n is the number of candidate predictor variables and I is the estimated event rate in the population [107]. An estimated event rate of 15% based on our pilot data was used; for a model with 15 effective candidate predictor variables (ie. dipstick proteinuria is counted three times to reflect inclusion of three indicator variables) the sample size required was 1000 women [116,126]. This sample size target was doubled to allow for subgroup analysis at the conclusion of the study after the finding of confounding by centre during the interim analysis.

3.2.6 Statistical methods

Development and validation of the miniPIERS model followed the general steps outlined in Chapter 2 of this thesis. Details of each step are described below.

Coding of predictors.

The relationship between each predictor variable and the combined adverse maternal outcome was first assessed by univariate logistic regression. Continuous variables were assessed for non-linearity, and were modeled as restricted cubic splines when appropriate [107]. Variables with a skewed distribution were log-transformed (natural log). Inclusion of the transformed variable in the final model was based on comparison to a model with the linear variable and selection of the model with the lowest Akaike information criterion (AIC) was automated during the model development process.

To avoid co-linearity, correlation between variables was determined and only the more clinically relevant variable of a pair of highly correlated variables was retained. When a high degree of correlation existed between two symptoms ($r>0.5$) they were re-coded as a combined indicator variable.

Model building.

Stepwise backward elimination was used to build the most parsimonious model with a stopping rule of $p<0{\cdot}20$. No interaction terms were included in the model as no interaction was hypothesized between candidate predictors prior to analysis.

We assessed the potential for confounding by study site by examining the bivariate association of study site with predictor variables and with outcome rate. Dummy (indicator) variables for study site were included in the model to eliminate confounding of the predictor-adverse outcome relationship by study site. To make the final model generalizable to all study settings, the coefficients for site variables were excluded from the calculation of predicted probability, and the model's intercept was adjusted using previously published methods for updating a prediction model for a new setting [112,127].

Assessing the model's performance.

Calibration ability of the model was assessed visually by plotting deciles of predicted probability of an adverse maternal outcome against the observed rate in each decile and fitting a smooth line [107,128]. Discrimination ability was evaluated based on the area under the receiver operating characteristic curve (AUC ROC) [129]. The sensitivity, specificity, positive predictive value, negative predictive value, and likelihood ratios (LR) of cut-offs for a positive test defined using the population within each risk group were calculated [130]. The following categories for interpretation of the likelihood ratios were used: informative (LR<0·1 or >10); moderately informative (LR 0·1-0·2 or 5-10); and non-informative (LR 0·2-5).

A risk stratification table was generated to assess the extent to which the model's predictions divided the population into clinically distinct risk categories [131].

Model validation.

Internal validation of the model was assessed using 500 iterations each of Efron's enhanced bootstrap method [114]. Details of this approach have been described previously [107,119]. The bootstrapping procedure involved (1) sampling with replacement from the original cohort to generate a bootstrap dataset of 2081 women; (2) redevelopment of the model including all model development steps; variable coding (transformations and categorizations), variable selection and parameter estimation in the bootstrapped sample; (3) estimation of the AUC ROC for the model in the bootstrap sample; (4) application of this new model to the original dataset and estimation of AUC ROC. Model optimism is then calculated as the average difference between model performance in the bootstrap sample and the original dataset after 500 iterations of this procedure. The choice was made to use 500 iterations because previous studies have shown no benefit is achieved when using a higher number of repetitions [108,116]. A final assessment of calibration was performed using the Hosmer-Lemeshow goodness-of-fit test.

A final assessment of model validity was performed by applying the miniPIERS model to the fullPIERS dataset and estimating the AUC ROC. Due to the marked difference in underlying rate of outcomes in the fullPIERS population (6.5% in fullPIERS vs 12.5% in miniPIERS), the model intercept (i.e. the baseline rate) was adjusted before estimating predictive performance [108,112]. This difference in outcome rate between the two cohorts is due to the difference in setting in

which the data were collected; as noted in the description of the cohorts above, fullPIERS was completed in high-income country facilities only.

Sensitivity analyses were performed to assess the generalizability of the model in various subsets of study data. In addition, sensitivity analyses were performed excluding the most common components of the adverse maternal outcome to ensure that model discriminatory ability was maintained. Generalizability of the model across study regions was further assessed based on the AUC ROC calculated for the model when applied to each region's subset of the total miniPIERS cohort.

All statistical analyses were performed using STATA v11·0 (StataCorp, College Station, TX, USA).

3.3 Results

From 1 July 2008 to 31 March 2012, 2133 women were recruited to the miniPIERS cohort. Fifty-two of these women were excluded prior to analysis after review of their medical record revealed that they were ineligible. Medical chart review was able to resolve all instances of missing predictor variables in the total cohort. Data relating to the remaining 2081 women were included in the model development and internal validation process. The breakdown in recruitment by site is provided as part of Table 3-9. Compared with women who did not have an adverse outcome, women who had an adverse outcome were more likely to be nuliparous, admitted earlier in gestation, admitted with a diagnosis of pre-eclampsia, had worse clinical measures in the first 24 hours of admission, and were more likely to have received

corticosteroids and magnesium sulphate but less likely to have been delivered by Caesarean section (Table 3-1).

Maternal adverse outcomes included two maternal deaths during the study. The most common morbidities that occurred were the need for blood transfusion (174 women (8·4%)), placental abruption (70 women (3·4%)), and pulmonary oedema (51 women (2·5%)) (Table 3-2). There were 32 (1·5%) women with one or more seizures of eclampsia after admission, of whom 31 received magnesium sulphate.

Initial variable inspection demonstrated that there was a strong correlation (r>0·5) between the symptoms of chest pain and dyspnoea, and headache and visual disturbances. Therefore, these symptoms were re-coded as combined indicator variables and entered accordingly into the multivariate model. As expected, systolic and diastolic blood pressure were highly correlated. Systolic blood pressure was selected for final model development because it is easier for minimally-trained health care providers to measure by radial artery palpation than detection of Korotokoff sounds and has been shown to be reflective of stroke risk in women with pre-eclampsia [76]. Systolic blood pressure measurements were log transformed for final model development as was gestational age at admission due to the highly skewed distribution of both variables (Figure 3-1). Non-linear transformations were also tested for systolic blood pressure and gestational age during final model development after plotting as restricted cubic splines revealed potential non-linear relationships with the outcome in univariate analysis (Figure 3-2).

Bivariate assessment of predictor variable effects with or without the centre variable did reveal evidence of confounding by center. A good example of this can be shown when assessing the effect of the dipstick variable on outcome, the dipstick effect was consistent across study sites and significant. Addition of the site variable had significant impact on the odds of outcome (Table 3-3) in all dipstick categories. In addition, detailed review of distribution and patterns of reporting in each site (Table 3-4) demonstrated differences between study sites in regards to populations included in the study. Due to these factors a site variable was included in the final model.

Table 3-5 presents results of the univariate and multivariate analysis of miniPIERS predictors. The final miniPIERS equation was: logit (logarithm of the odds)(pi)= -5.77 + [-2.98×10^{-1} x indicator for multiparity] + [(-1.07)x log gestational age at admission] + [1·34 x log systolic blood pressure] + [($-2{\cdot}18 \times 10^{-1}$)x indicator for 2+ dipstick proteinuria] + [($4{\cdot}24 \times 10^{-1}$) x indicator for 3+ dipstick proteinuria] + [(5.12×10^{-1}) x indicator for 4+ dipstick proteinuria] + [1·18 x indicator for occurrence of vaginal bleeding with abdominal pain] + [(4.22×10^{-1}) x indicator for headache and/or visual changes] + [8.47×10^{-1} x indicator for chest pain and/or dyspnoea].

The model appeared well-calibrated, as shown in the calibration plot (Figure 3-3). In all deciles except for the highest the 95% confidence interval around the observed outcome rate crossed the diagonal fitted line. The AUC ROC for this model was 0·768 (95% CI 0·735 – 0·801) (Figure 3-4) with an average optimism estimated to be 0.037. Using a cut-off of predicted probability of 25% to define a positive test resulted in a likelihood ratio of 5.09 [4.12 - 6.29] and classified

women with 85.5% accuracy (sensitivity 41.4%; specificity 91.9%). The stratification capacity of the model was good, as shown by the 784 (37.7%) and 256 (12.3%) women in the lowest and highest risk groups, respectively (Table 3-6).

Data from 1300 women in the fullPIERS cohort were used for external validation of the developed miniPIERS model. Table 3-7 presents the results of a comparison of demographics and clinical characteristics of women in fullPIERS compared to miniPIERS. The cohorts differed significantly with respect to demographics, interventions and pregnancy outcomes. When the miniPIERS model was applied to the fullPIERS dataset the AUC ROC was 0.713 [95% CI 0.658 – 0.768] after adjusting the model intercept to account for differences in the outcome rate between the fullPIERS and miniPIERS populations (Figure 3-5).

The results of several sensitivity analyses done using the miniPIERS cohort are presented in Table 3-8. In all subsets, model performance was maintained. Of note, when the cohort was restricted to only those women admitted with a diagnosis of pre-eclampsia (defined as hypertension and proteinuria) the AUC ROC was 0.77 [0.73 - 0.81]. In addition, when including the whole cohort but restricting the definition of the adverse outcome to include only maternal death, eclampsia, stroke, cortical blindness or retinal detachment the AUC ROC was 0.81 [0.75 - 0.87]. The model performance did not appear to differ significantly between study regions, although the confidence interval around the estimate of the AUC ROC in small study sites was wide (see Table 3-9).

Table 3-8 also presents sensitivity analyses performed using the fullPIERS cohort. Due to the smaller number of events in this cohort, not all analyses could be meaningfully repeated but where performed, model performance appeared to be maintained.

3.4 Discussion

3.4.1 Main findings

Using data from a prospectively collected cohort of 2081 women with HDP admitted to a hospital in five LMICs, we have developed and internally validated the miniPIERS model. The final miniPIERS model includes only demographics, symptoms and signs that can be measured in primary health care facilities in low-resourced settings. Data for the study were collected by nurses and research staff with basic training to ensure the feasibility of replication of the measurements by comparable workers. For example, gestational age can be estimated from clinical information when ultrasound is unavailable, symptoms can be ascertained with simple questions, systolic blood pressure can be estimated easily using the radial pulse, and dipstick proteinuria can be estimated by assessing the opacity of boiled urine when dipsticks are not available [132]. By limiting the model to these simple measures, the miniPIERS model has potential for use by mid-level health workers in low-resourced settings. To add to the ease of use of this model, miniPIERS is being converted to a mobile health application that will be useable on any mobile device so that health care workers are not required to calculate risk directly, as described in the next chapter.

Overall, the miniPIERS model performed well based on accuracy and discrimination ability (i.e. the AUC ROC). There was a slight underestimation of risk in the highest decile of predicted probability, but because the model was designed to be used as a categorical decision rule, this error in calibration is not thought to be clinically relevant. This model attains similar stratification, calibration, and classification accuracy as other established risk scores used in adult and reproductive medicine [104,133]. To our knowledge, the miniPIERS model is the only clinical prediction model developed and validated for use with pregnant women in LMICs.

The miniPIERS model was used to designate women as being high-risk if their predicted probability of adverse outcome was ≥25%. This threshold was chosen based on the associated 10% false positive rate and approximately 50% sensitivity. A full description of the process of determining the optimal threshold at which to classify women as high-risk is described in the following chapter. The likelihood ratio associated with this threshold showed potential utility as a rule-in test for adverse maternal outcome. By improving the ability of care providers to identify women at high risk of adverse outcomes, our specific aim was to reduce triage delays for women with any HDP in LMICs. What may be most useful is to set one threshold of predicted probability of adverse outcome, such as >15%, to initiate increased surveillance and use the higher threshold of ≥25% to initiate transport to a facility where emergency obstetric care is available. The positive predictive value of the 25% threshold was approximately 40% in all datasets with a corresponding 85% classification accuracy. These modest results highlight the fact that demographics, symptoms and signs alone will not identify all women with severe

disease but still have the potential to significantly improve care in resource limited areas and community settings where no or minimal monitoring of women with the HDP currently occurs.

3.4.2 Strengths and limitations

There are several limitations to this study. The first is the use of a combined adverse maternal outcome comprised of events of unequal severity. The Delphi consensus group determined that all components of the outcome were important enough on their own to warrant avoidance. The sensitivity analyses performed using a restricted definition of the adverse maternal outcome demonstrated that the model maintained its performance even when the more common and less-severe outcomes were excluded. A second limitation of the study is the use of broad inclusion criteria that included women with any HDP. This decision was made to make the model maximally useful for women who present with HDP, and for whom the exact diagnosis may not (or cannot) be determined at the time of clinical presentation. Reassuringly, when we restricted the cohort to only those women who were admitted with classically defined pre-eclampsia (hypertension and proteinuria), model performance was maintained.

A third limitation is the use of a backward elimination method for final variable selection in the model. Automated variable selection methods for model development have been shown to be sensitive to minor changes in the data and are not easily reproducible [134]. Ultimately, we felt that creating a simpler model with only those few variables that were most predictive of the outcome was important to make application of the model by minimally trained care providers easier.

A fourth limitation is the use of the fullPIERS dataset for external validation of the model. Although the data were collected for both fullPIERS and miniPIERS using the same definitions and protocols, the populations between the two studies differed significantly, as did the care received. Ideally the model should be validated in another cohort of data from low-resourced settings collected by mid-level care providers as part of routine care. This is planned and would address the possible concern for a reduction in model performance should these health workers be unable to maintain the level of measurement accuracy achieved in the facility data we have used for this study. In the interim, it was reassuring that there was consistency of results between fullPIERS and miniPIERS models. miniPIERS model performance was maintained in the fullPIERS cohort and more importantly coefficients were similar in overlapping predictors between the fullPIERS and miniPIERS models. This gives us confidence that this is a well-defined and stable model.

A final limitation is the inclusion of clinically-defined gestational age within the miniPIERS model, usually based on last menstrual period dates. As in fullPIERS, increasing gestational age was associated with diminishing risk [119]. This inverse relation was maintained in this study despite the inaccuracy inherent in clinically-based gestational age assessment. Despite these limitations we were able to achieve accurate predictions from the miniPIERS model.

A major strength of this study is the high quality of data collected in a standardized manner. We were able to ensure that complete data were collected in five different LMICs through careful study monitoring and training of research staff. A second strength of this study is the

generalizability of the resulting model. By combining high quality data from multiple international sites we are able to generate a model that should be applicable to any LMIC setting. The generalizability of the model is further supported by the results of the region-specific analysis of model performance. It is likely that we would have had greater predictive power had we developed the model using a more homogeneous population from one geographic region, but this would have resulted in a less generalizable model. By trading some predictive ability for generalizability, we believe we will have achieved greater impact on global public health. A final strength of the study is the use of clinically important timeframes for assessment and prediction. The miniPIERS model predicted adverse maternal outcomes occurring within 48 hours of assessment using data from within 24 hours of assessment; such timeframes represent clinically useful time periods in which transportation or disease-modifying interventions such as magnesium sulphate, antihypertensive agents, and delivery can be initiated.

3.4.3 Interpretation and conclusion

The miniPIERS model is the first clinical prediction model developed and validated specifically for use in low-resourced settings by minimally trained health workers. This model represents a significant step forward in our ability to provide evidence-based management of women with HDP in these settings. The potential impact of use of this model as a part of routine antenatal care is significant, given the high burden of outcomes as a result of the HDP in low-resourced settings. Nevertheless, as with any prediction model, its ultimate value will only be

demonstrated with an implementation project that is able to demonstrate that its potential can be translated to real health systems change and clinical improvements; such a project, called the Community Level Interventions for Pre-eclampsia (CLIP) study (clinicaltrials.gov ID NCT01911494), is presently underway. For more information on the CLIP study, please see http://pre-empt.cfri.ca/OBJECTIVES/CLIPTrial.aspx). Until that study is complete, the miniPIERS model can be used as a basis of a community education programme to increase women's, families', and community-based health workers' knowledge of warning symptoms and signs associated with the HDP.

Table 3-1: Demographics of women in the total cohort comparing women with and without adverse maternal outcomes (N=2081). Results for continuous variables presented as mean (+/- sd) when data normally distributed or median [interquartile range] for skewed data.

Characteristic	Women with adverse outcomes (n= 401 women)	Women without adverse outcomes (n= 1680 women)	P*
Demographics (within 48h of eligibility)			
Maternal age at EDD (years) mean (+/- sd)	27·9 (+/- 5·9)	28·5 (+/- 6·2)	0·17
Parity ≥ 1 n(%)	183 (45·6%)	939 (55·9%)	<0·01
Gestational age at eligibility (weeks) median [interquartile range]	35·3 [30·7, 38·1]	37·1 [34·1, 38·8]	<0·01
Multiple pregnancy n(%)	17 (4·2%)	57 (3·4%)	0·41
Smoking in this pregnancy n(%)	25 (6·2%)	72 (4·3%)	0·08
HDP description			
Pre-eclampsia n(%)	320 (79·8%)	1016 (60·5%)	<0·01
Other HDP n(%)	81 (20·2%)	664 (39·5%)	<0·01
Clinical measures (within 24h of eligibility)			

Characteristic	Women with adverse outcomes (n= 401 women)	Women without adverse outcomes (n= 1680 women)	P*
Systolic BP (mmHg) median [interquartile range]	170 [150, 186]	150 [140, 170]	<0·01
Diastolic BP (mmHg) median [interquartile range]	110 [100, 120]	100 [90, 110]	<0·01
Worst dipstick proteinuria median [interquartile range]	2+ [1+, 3+]	1+ [trace, 3+]	<0·01
Number of symptoms median [interquartile range]	1 [0, 2]	0 [0, 1]	<0·01
Interventions at any time during admission			
Corticosteroid administration n (%)	180 (44·9%)	525 (31·3%)	<0·01
Antihypertensive medications administered n (%)	386 (96·3%)	1560 (92·9%)	0·13
$MgSO_4$ administered n (%)	271 (67·6%)	677 (40·3%)	<0·01
Pregnancy outcomes			
Admission-to-delivery interval (all cases) (d) median [interquartile range]	1 [1, 4]	1 [1, 5]	0·02
GA on delivery (weeks) median [interquartile range]	35·7 [31·7, 38·3]	37·6 [35·3, 39·1]	<0·01

Characteristic	Women with adverse outcomes (n= 401 women)	Women without adverse outcomes (n= 1680 women)	P*
Delivery at <34+0 weeks GA n (%)	160 (39.9%)	290 (17.3%)	<0·01
Caesarean Delivery n (%)	110 (27.4%)	625 (37.2%)	<0·01
Birth weight (g) median [interquartile range]	2100 [1303, 2800]	2700 [2000, 3150]	<0·01
Birth weight <3rd percentile (N babies) n (%)	64 (16·0%)	284 (16·9%)	0·66
Intrauterine fetal death (≥20+0 wk and/or ≥500g) n (%)	54 (13·5%)	94 (5·6%)	<0·01
Neonatal death (before discharge) n (%)	26 (6·5%)	42 (2·5%)	<0·01

*p values calculated using chi-squared test for categorical variables and student's t-test or Mann-Whitney U for continuous variables

EDD = estimated date of delivery; HDP = hypertensive disorder of pregnancy; BP = blood pressure; GA = gestational age

Table 3-2: Maternal adverse outcomes occurring in the total miniPIERS cohort, outcome counts not mutually exclusive when listed within 48 hours or at any time during admission.

	Total Cohort (N=2081)	
One or more of maternal morbidity or mortality:	**within 48h**	**any time**
TOTAL n(%)	**261 (12·5%)**	**401 (19·3%)**
Maternal death	1	2
Central nervous system		
Eclampsia (≥1)	24	32
Glasgow coma score <13	8	11
Stroke or reversible ischaemic neurological deficit	3	4
Cortical blindness or retinal detachment	4	5
Posterior reversible encephalopathy	0	1
Cardiorespiratory		
Positive inotropic support	2	3
Infusion of a 3rd parenteral antihypertensive	8	9
Myocardial ischaemia/infarction	2	4
SpO_2 <90%	9	22
≥50% FiO_2 for >1hr	5	7
Intubation (other than for Caesarean section)	14	25
Pulmonary oedema	37	51
Haematological		
Transfusion of any blood product	129	174
Platelets <50 x 10^9/L with no transfusion	15	19
Hepatic		
Dysfunction	7	9

	Total Cohort (N=2081)	
One or more of maternal morbidity or mortality:	**within 48h**	**any time**
Haematoma/rupture	0	0
Renal		
Acute renal insufficiency	21	28
Dialysis	1	2
Placental outcomes		
Placental abruption	39	70
PPH requiring hysterectomy	39	50
Other adverse events		
Severe ascites	26	46
Other**	3	8

**includes 5 cases of pulmonary embolism, 2 cardiac arrests, 1 ruptured uterus

SpO_2 = blood oxygen saturation; FiO_2 = fractional inspired oxygen; PPH = postpartum haemmorhage.

Table 3-3: Effect estimate for dipstick on occurrence of adverse maternal outcome before and after adjustment by study centre.

	Unadjusted OR (95% CI)	**Adjusted OR (95% CI)**	**% change in OR**
Negative or trace	Reference	Reference	Reference
1+	1.36 (0.89, 2.01)	1.55 (0.99, 2.42)	13.9%
2+	2.04 (1.41, 2.96)	2.57 (1.65, 4.00)	25.9%
>3+	1.88 (1.20, 2.94)	4.50 (2.68, 7.55)	139%

Table 3-4: Outcome rates, demographics and clinical characteristics by centre. Six centres are presented because the Pakistan hospitals where grouped according to location, Karachi vs. Hyderabad, at the site investigators request.

Variable	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6
Outcome in 48hrs	6.8%	3.9%	5.2%	19.2%	7.0%	26.0%
Maternal Age at EDD	29.2 (+/-5.5)	29.4 (+/-6.9)	26.4 (+/-5.7)	27.1 (+/-6.8)	28.7 (+/-7.5)	29.4 (+/-5.6)
GA on eligibility	37.5 [35.1, 38.9]	36.2 [32.8, 37.9]	36.8 [33.6, 39.4]	31.9 [28.4, 35.8]	37.1 [34.2, 38.9]	37.2 [35.6, 38.7]
Parity ≥ 1	52.0%	53.3%	56.1%	54.4%	56.1%	53.6%
sBP	140 [140, 155]	150 [140, 151]	170 [150, 180]	176 [162, 190]	160 [140, 170]	150 [140, 170]
dBP	90 [90, 100]	100 [90, 100]	110 [100, 125]	112 [110, 120]	100 [90, 110]	110 [100, 110]
Dipstick	Trace [neg, 1+]	1+ [trace, 2+]	3+ [1+, 4+]	3+ [2+, 3+]	Trace [neg, 1+]	2+ [2+, 3+]
Number of symptoms	0 [0, 1]	0 [0, 0]	0 [0, 1]	1 [0, 1]	0 [0, 1]	2 [1, 3]
MgSO4 use	16.1%	27.6%	75.7%	78.2%	31.0%	47.3%
Anti-htn use	88.8%	69.3%	99.7%	99.4%	98.4%	95.2%

Variable	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6
Adm – delivery (d)	1 [0, 2]	7 [2, 18]	3 [2, 5]	2 [1, 8]	2 [1, 6]	1 [1, 2]
Birthweight (g)	2800 [2300, 3200]	2925 [2300, 3450]	2500 [1800, 3000]	1690 [1142, 2360]	2695 [2275, 3425]	2500 [2000, 3000]
Stillbirth	1.6%	3.1%	11.6%	6.3%	1.1%	15.4%
Neonatal death	1.1%	0.7%	6.1%	4.0%	1.6%	4.8%

Anti-htn = antihypertensive; EDD = estimated date of delivery; GA = gestational age; HDP = hypertensive disorder of pregnancy; sBP = systolic blood pressure; dBP = diastolic blood pressure; Adm= admission

Table 3-5: Univariate and multivariate analysis of candidate predictors in the miniPIERS cohort. Variables presented as part of the multivariate analysis are those that were retained after model development and backward selection.

Candidate Predictor	Univariate OR [95% CI]	Multivariate OR [95% CI]
Demographics		
Maternal age (years)	0.99 [0.97, 1.01]	n/a
Gestational age at admission (weeks)	0.95 [0.92, 0.98]	0.34 [0.11, 1.11]*
Parity (multip vs. primip)	0.73 [0.57, 0.95]	0.74 [0.56, 0.99]
Signs		
Systolic blood pressure (mmHg)	1.02 [1.01, 1.02]	3.89 [1.19, 12.66]*
Diastolic blood pressure (mmHg)	1.03 [1.02, 1.03]	n/a
Dipstick proteinuria		
2+	1.44 [0.99, 2.09]	0.80 [0.51, 1.27]
3+	2.88 [2.07, 4.00]	1.53 [0.99, 2.37]
4+	3.23 [2.18, 4.85]	1.67 [0.97, 2.88]
Symptoms		
Headache	3.42 [2.58, 4.52]	1.53 [1.07, 2.17]
Visual disturbances	2.63 [2.00, 3.45]	
Chest pain	6.42 [3.62, 11.37]	2.33 [1.38, 3.94]
Dyspnoea	6.35 [4.08, 9.89]	
Epigastric/ Right upper quadrant pain	3.93 [2.96, 5.21]	n/a
Nausea/ vomiting	3.40 [2.53, 4.57]	n/a
Abdominal pain with vaginal bleeding	6.03 [4.25, 8.57]	3.24 [2.13, 4.94]

*log transformed

Table 3-6: Risk stratification table to assess the miniPIERS prediction model Upper limit of predicted probability range used to define a positive test for sensitivity (Sens), specificity (Spec), positive predictive value (PPV) and negative predictive value (NPV)

Predicted probability	# event/ # in range	Sens %	Spec %	PPV %	NPV %	LR [95% CI]*
0 – 5·5%	33/784	-	-	-	-	0.31 [0.22, 0.42]
5·6 -8·0%	18/286	87.4	41.3	17.6	95.8	0.47 [0.29, 0.74]
8·1 – 15·0%	46/456	80.5	56.0	20.8	95.2	0.78 [0.59, 1.03]
15.1 – 24.9%	56/299	62.8	56.6	29.5	93.6	1.61 [1.24, 2.08]
≥ 25%	108/256	41.4	91.9	42.2	91.6	5.09 [4.12, 6.29]

*likelihood ratio (LR) for each category calculated using the method described by Deeks et al [130]

Table 3-7: Demographics of women in the total cohort comparing women with and without adverse maternal outcomes (N=2081). Results for continuous variables presented as mean (+/- sd) when data normally distributed or median [interquartile range] for skewed data.

Characteristic	miniPIERS cohort (n= 2081 women)	fullPIERS cohort (n= 1300 women)	P*
Demographics (within 48h of eligibility)			
Maternal age at EDD (years) mean (+/- sd)	28.4 (+/- 6.2)	31.7 (+/- 6.0)	<0.01
Parity ≥ 1 n(%)	1122 (53.9%)	403 (31.0%)	<0.01
Gestational age at eligibility (weeks) median [interquartile range]	36.8 [33.5, 38.7]	37.0 [34.1, 38.9]	0.04
Pre-eclampsia description			
Pre-eclampsia n(%)	1336 (64.2%)	1020 (78.5%)	<0.01
Other HDP n(%)	745 (35.8%)	280 (21.5%)	<0.01
Clinical measures (within 24h of eligibility)			
Systolic BP median [interquartile range]	160 [140, 170]	166 [155, 180]	<0.01
Diastolic BP median [interquartile range]	100 [95, 110]	104 [98, 110]	0.22

Characteristic	miniPIERS cohort (n= 2081 women)	fullPIERS cohort (n= 1300 women)	P*
Worst dipstick proteinuria median [interquartile range]	2+ [trace, 3+]	1+ [trace, 3+]	0.01
Number of symptoms median [interquartile range]	1 [0, 1]	1 [0, 2]	<0.01
Interventions at any time during admission			
Corticosteroid administration n (%)	705 (33.9%)	337 (25.9%)	<0.01
Antihypertensive medications administered n (%)	1946 (93.5%)	836 (64.3%)	<0.01
$MgSO_4$ administered n (%)	948 (45.5%)	370 (28.5%)	<0.01
Pregnancy outcomes			
Admission-to-delivery interval (all cases) (d) median [interquartike range]	1 [1, 4]	1 [1, 4]	0.24
GA on delivery (weeks) median [interquartile range]	37.3 [34.6, 39.0]	37.6 [35.3, 39.1]	0.16
Delivery at <34+0 weeks GA n (%)	450 (21.6%)	319 (24.5%)	0.04
Adverse maternal outcome (within 48hrs of admission) n (%)	261 (12.5%)	84 (6.5%)	<0.01

Characteristic	miniPIERS cohort (n= 2081 women)	fullPIERS cohort (n= 1300 women)	P*
Birth weight (g) median [interquartile range]	2600 [1900, 3090]	2836 [2105, 3365]	<0.01
Intrauterine fetal death (≥20+0 wk and/or ≥500g) n (%)	148 (7.1%)	15 (1.2%)	<0.01
Neonatal death (before discharge) n (%)	68 (3.3%)	14 (1.1%)	<0.01

*p values calculated using chi-squared test for categorical variables and student's t-test or Mann-Whitney U for continuous variables

EDD = estimated date of delivery; HDP = hypertensive disorder of pregnancy; BP = blood pressure; GA = gestational age

Table 3-8: Results of sensitivity analysis using the miniPIERS model to predict adverse maternal outcome in subsets of the data or to predict restricted definition of the combined adverse outcome, as described, in the miniPIERS and fullPIERS cohorts.

Cohort description	Outcome incidence in miniPIERS cohort (n/N)	AUC ROC [95% CI]	Outcome incidence in fullPIERS cohort (n/N)	AUC ROC [95% CI]
Including only women admitted with diagnosis of pre-eclampsia*	200/1336	0.77 [0.73, 0.81]	73/1028	0.72 [0.65, 0.79]
Including all but blood transfusion as adverse maternal outcome	174/2081	0.76 [0.72, 0.80]	68/1300	0.76 [0.73, 0.78]
Including all but PPH and placental abruption as adverse maternal outcome	240/2081	0.78 [0.74, 0.81]	n/a	n/a
Including maternal mortality, eclampsia, stroke, retinal detachment or cortical blindness occurring at any time after admission only	38/2081	0.81 [0.75, 0.87]	n/a	n/a
Including only women admitted ≤34+6 weeks GA	94/578	0.76 [0.70, 0.82]	n/a	n/a
Including only women admitted >34+6 weeks	167/1503	0.77 [0.72, 0.81]	49/973	0.73 [0.64, 0.82]

Cohort description	Outcome incidence in miniPIERS cohort (n/N)	AUC ROC [95% CI]	Outcome incidence in fullPIERS cohort (n/N)	AUC ROC [95% CI]
Including only women admitted ≥37+0 weeks GA	108/997	0.78 [0.73, 0.83]	n/a	n/a

*other hypertensive disorders excluded: chronic hypertension, gestational hypertension without proteinuria or other adverse conditions, partial HELLP

GA: gestational age; AUC ROC: area under the receiver operating characteristic curve

Table 3-9: Performance of the model in each study site region as a predictor of combined adverse maternal outcome occurring within 48hrs of admission

Region	Contribution of cases to total miniPIERS cohort (%)	Outcome incidence in cohort used (n/N)	AUC ROC (95% CI)
Brazil	9.0	13/187	0.69 [0.52, 0.83]
Fiji	6.1	5/127	0.72 [0.49, 0.95]
Pakistan	50.7	157/1056	0.76 [0.71, 0.80]
South Africa	16.8	67/349	0.76 [0.70, 0.82]
Uganda	17.4	19/362	0.66 [0.51, 0.80]

AUC ROC: area under the receiver operating characteristic curve

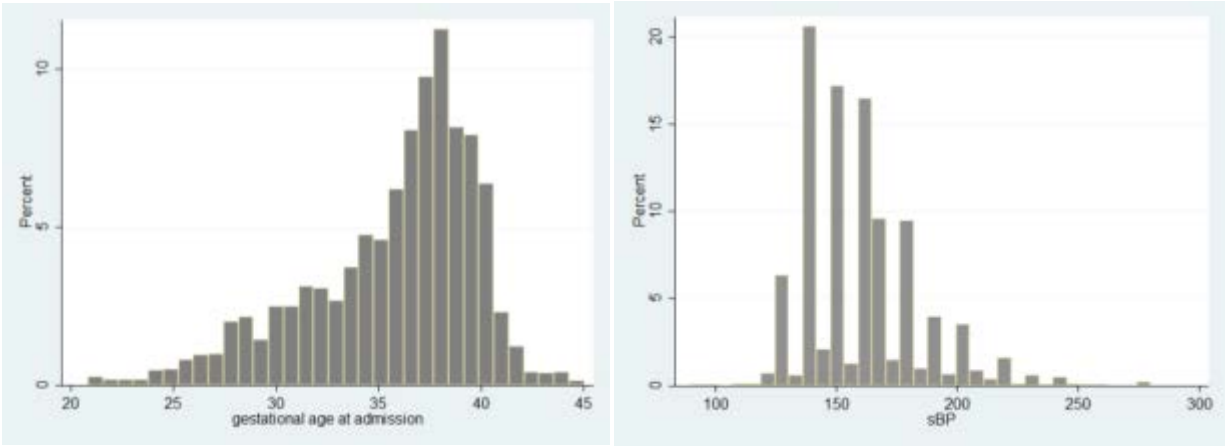


Figure 3-1: Histograms showing frequency distribution of (a) gestational age on admission and (b) systolic blood pressure measured within 24 hours of admission.

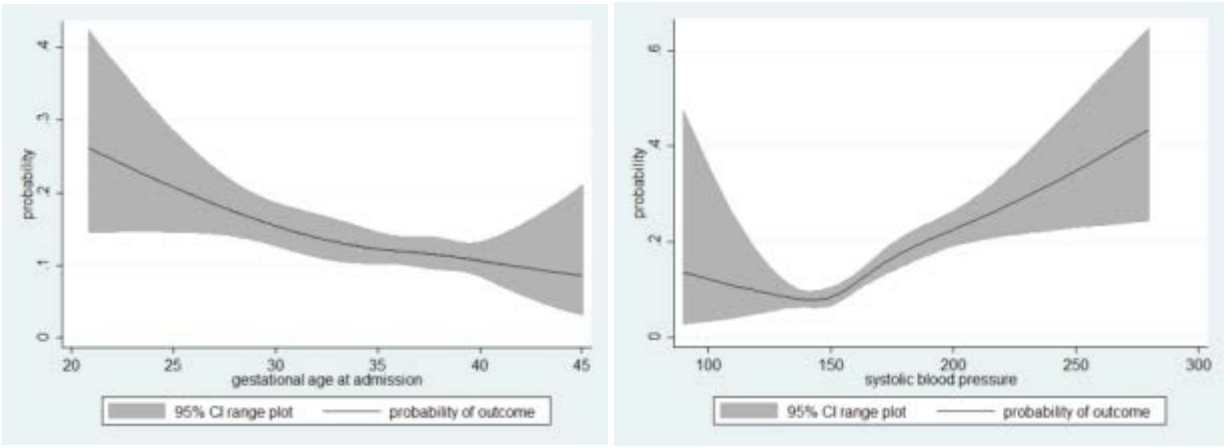


Figure 3-2: Restricted cubic spline with 4 knots exploring potential non-linear relationships between (a) gestational age on admission and (b) systolic blood pressure measured within 24 hours of admission with occurrence of adverse maternal outcome within 48 hours of admission.

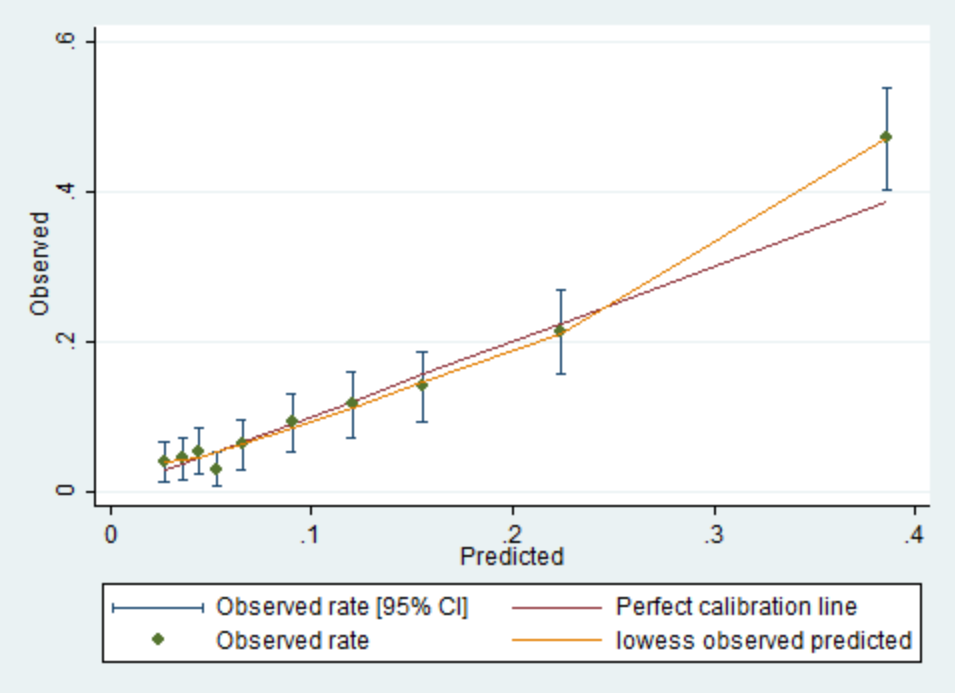


Figure 3-3: Calibration plot showing fit of the model in the miniPIERS cohort by comparing observed outcome rates against 10 deciles of predicted probability as defined by the model.

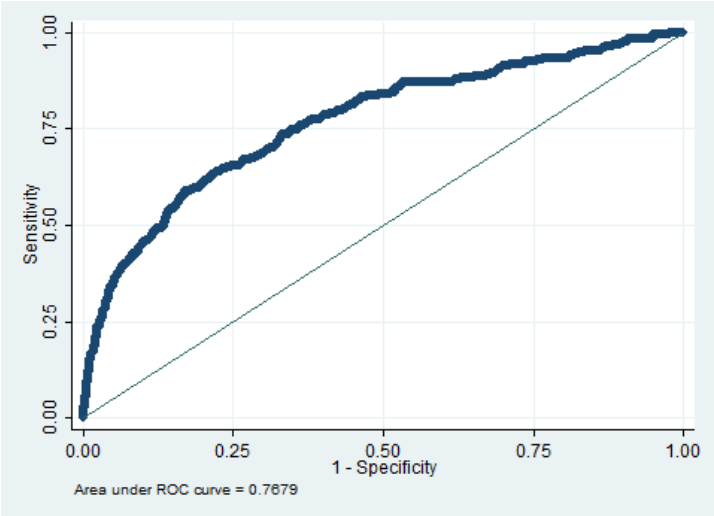


Figure 3-4: Receiver operating characteristic curve for miniPIERS model applied to the development dataset (AUC 0.768; 95% CI 0.735 – 0.801)

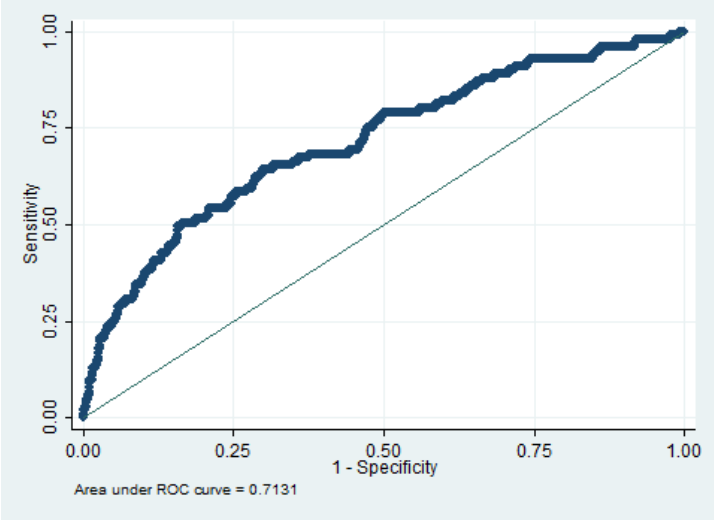


Figure 3-5: Receiver operating characteristic curve for miniPIERS model applied to the fullPIERS dataset (AUC 0.713; 95% CI 0.658 – 0.768)

Chapter 4: Design and assessment of accuracy of a mobile health decision aid based on the miniPIERS model for use in low-resourced settings

4.1 Introduction

The miniPIERS model, as described in the previous chapter, was developed and validated for use in low-resourced settings by minimally trained health workers [135]. The model itself includes variables that are simple to collect at low cost and require only basic training and equipment to measure, such as an automated blood pressure monitor. Application of the miniPIERS model into a community setting requires translation of this complex equation into a format that can allow easy calculation of risk and that will guide the health worker through an appropriate clinical response to that risk. Options for application of a clinical prediction model that have been previously established include conversion of the model into a score card, a nomogram, or development of an online calculator [73,111,127].

Score cards have routinely been used in clinical practice for application of risk models. Examples include the Framingham risk score [136] and the Acute Physiology and Chronic Health Evaluation (APACHE) score [104]. These score cards have been shown to be effective in high-resourced settings when used by well-trained physicians and are simple to understand but require dichotomization of the continuous predictors included in the models in order to establish risk categories[137]. Dichotomization of continuous variables within a prediction model reduces accuracy of the model and should be avoided when possible [138,139](merged). Nomograms are similar to score cards in that they allow creation of risk groups from a continuous risk scale but

also present a graphical display of the prediction model [140](merged). Some dichotomization of variables is still required for development of the nomogram but they have been shown to better maintain accuracy of predictions when properly interpreted [140](merged). Criticisms of the nomogram are that it is difficult to understand and requires a high-degree of literacy [137]. A good example of a nomogram that has been developed for use in reproductive medicine can be found in the application of a model for assessing likelihood of success of vaginal birth after Caesarean delivery [141]. Computer programs, either online or using mobile phone applications that allow application of risk models are now becoming more prominent, given the inherent flaw in dichotomization of predictors during development of score cards and nomograms and the difficulty in interpreting the nomogram results.

Given the availability of mobile phones in the LMIC settings in which we hope to apply the miniPIERS model [70], and the initial success of mHealth programs for improved maternal health in low-resourced settings [71], in collaboration with the miniPIERS working group members we determined that conversion of the miniPIERS model into a mobile phone application would be the optimal method of application of this model into community-based antenatal care. A mobile phone application would allow the added benefit of inclusion of other important risk markers and decision points defined based on the needs of the local health workers and established pre-eclampsia management guidelines [7]. Incorporation of multiple decision points, including the miniPIERS risk assessment requires development of an interactive and multi-leveled decision algorithm on which the mHealth tool could be based. This decision algorithm should be designed to reflect the needs of the target users (community health workers), and

should reflect the complex and multi-faceted nature of the hypertensive disorders of pregnancy (HDP) and the various clinical requirements for assessment and management of the disorder in the field.

In this chapter, the process of developing the decision algorithm on which the mHealth application would be based and assessing its accuracy in the miniPIERS cohort is described. Technical development of this mobile phone application has been described elsewhere [142]. The final miniPIERS phone application is called the PIERS on the Move (POM) mHealth application.

4.2 Methods

Development of the decision algorithm involved multiple stages of review and feedback with relevant stakeholders including obstetricians, internists, midwives, nurses and research staff working within the PRE-EMPT (Pre-eclampsia and Eclampsia Monitoring, Prevention and Treatment) initiative. The first stage of this process required converting the miniPIERS model into a decision rule by defining the threshold of predicted probability that would be used to designate women as high-risk and who would require urgent referral and treatment. Once this was complete, integration of other clinical decision points into a broader decision algorithm and incorporation of the WHO recommendations for management of women with pre-eclampsia or eclampsia was completed over several iterations of discussion with the PRE-EMPT team members.

4.2.1 Creating a decision rule based on the miniPIERS clinical risk prediction model

To determine the optimal cut-point of predicted probability at which to identify women as high-risk of adverse maternal outcome requiring intervention, a survey of the study working group members based on the concept of a discreet choice exercise was performed. A discreet choice exercise is a type of questionnaire designed to test alternative priorities in a clinical scenario and the effect of associated attributes of the decision on the priority set. The scenario around which the survey was designed to test was clinician preference for sensitivity (ability to identify all high risk cases correctly) vs. specificity (ability to identify low-risk cases correctly) when applying the miniPIERS model as a population screening tool at the community level. Design of the DCE followed standard methods for use of a DCE in health care and was as follows: (1) characterizing the decision and decision makers; (2) identifying relevant attributes that may influence choice; (3) development of choice sets; (4) applying the questionnaire to relevant decision makers [143].

In this case the decision of interest was defined as referral to a higher level of care (with or without community-based treatment) vs. continued antenatal monitoring at home or in a clinic. The decision makers would be the community health workers. Attributes of interest that could be considered to have an effect on the decision were defined based on review of the literature and existing guidelines for the management of women with hypertensive disorders of pregnancy [6,7,120]. Through discussion with the study coordinating team at UBC (BAP, PvD and LAM) it was determined that important attributes to consider would be those that vary the

perceived severity of the HDP and would be evaluated in the community setting, specifically, gestational age and blood pressure. With these parameters defined, choice pairs were set as shown in questionnaire presented in Appendix C. Not all possible combinations of attributes were used as we were simply trying to initiate discussion and build consensus within the working group around how best to use the miniPIERS model in a community setting. In a formal DCE all possible choice pairs would be tested and a choice model built; this was not done for the current study.

The final questionnaire was provided to all miniPIERS study working group members during the annual PRE-EMPT initiative meeting on November 9, 2011 and responses analyzed to determine proportion showing preference for true positive vs. false positive choices. After completing the questionnaire, respondents were presented with an interim version of the miniPIERS risk stratification table (Table 3-6) and asked to come to a consensus regarding the cut-point for use as a trigger for community treatment and referral based on the sensitivity, specificity, positive and negative likelihood ratios calculated for various possible levels of predicted probability.

4.2.2 Development of the decision algorithm and defining recommendations for community-based care of women with HDP

Development of the final decision algorithm on which to base the design of the miniPIERS mobile phone application involved an iterative review process with members of the PRE-EMPT working groups. In the first stage of this process, working group members from each

participating country were asked to define the usual components (demographics, symptoms and signs evaluated) of an antenatal clinic visit and highlight all clinical decision points that were felt to be relevant to the community assessment setting using a decision tree framework provided (Appendix D). This group exercise was performed during the PRE-EMPT annual meeting on November 9, 2011.

The decision tree framework provided included decision nodes listing all clinical measures included in the miniPIERS model along with treatment end-points that were pre-defined based on WHO guidelines for the management of women with HDP [7]. Each group of clinicians from each PRE-EMPT participating site (China, Brazil, Nigeria, Pakistan, South Africa, and Uganda) were asked to define thresholds associated with the decision node at which intervention would be required. For example, they identified the blood pressure value at which they would initiate community-based treatment with antihypertensives and the blood pressure threshold for referral to a higher-level health care facility.

Results of this exercise were recorded and summarized at the end of the group session and used to propose an initial list of decision points or triggers for referral and treatment to include in the algorithm. This initial list was then converted into a questionnaire (Appendix E) and provided to all PRE-EMPT group members at the annual meeting on November 7, 2012. On this questionnaire, the respondents were asked simple yes or no questions to determine whether they agreed with the identified treatment and referral triggers and if not, they were asked to provide additional comments and alternative suggestions. Finally, results of this questionnaire

were combined with the previous group exercise results to create an initial draft of the decision model. This was then reviewed one final time by all respondents who had participated at any stage of development of the decision model and final group consensus was obtained through iterative online discussion. A full version of the decision algorithm meant for use in facility and a simplified version for community triage were generated.

4.2.3 Assessment of the accuracy of the decision algorithm

Study Population

Data for this study were derived from the miniPIERS cohort as described in Chapter 3, section 3.2. Data collected during the miniPIERS study included demographics (maternal age, parity, and gestational age on admission); symptoms (headache, visual disturbances, chest pain/ dyspnoea, right upper quadrant pain or epigastric pain, nausea, vomiting, and vaginal bleeding with abdominal pain); and signs (blood pressure, and dipstick proteinuria). All data collection procedures were as described in Chapter 3, section 3.2.1. The main outcome measure was occurrence of one or more component of the combined adverse maternal outcome, as previously described, occurring at any point after admission to hospital. Women who met the primary outcome criteria were considered the true high-risk population for the purpose of analysis.

Statistical Analysis

Accuracy was measured as the proportion of women who would go on to experience one or more component of the miniPIERS adverse maternal outcome who were identified as high-risk

using the defined clinical decision points in the POM algorithm, combined with the proportion of women who did not experience an outcome who were identified as not being high-risk. Only the clinical decision points or triggers, including severe systolic blood pressure (≥160 mmHg), significant proteinuria (≥4+) and high miniPIERS probability (≥25%) were assessed in this analysis. Information on fetal movements was not collected during the miniPIERS study and the severe outcomes used as triggers in the decision algorithm were defined as study end-points, as they are events that we hope to avoid if the algorithm is effective. Expected rates of referral when the POM algorithm was applied to the cohort were determined as the total number of women identified based on one or more clinical decision point. Sensitivity and specificity of each decision point along with the combined algorithm were also calculated.

All statistical analyses were performed using STATA v11·0 (StataCorp, College Station, TX, USA).

4.3 Results

4.3.1 Defining the miniPIERS decision rule

During the annual PRE-EMPT meeting in Vancouver, BC on 9th of November 2011, 15 clinicians from Brazil (1); China (2); Canada (3); Pakistan (3); South Africa (2); Uganda (3); and the United States (1) completed the discrete choice survey. Responses to the survey showed that regardless of opportunity for treatment in the community, clinicians valued high sensitivity over specificity in the screening tool. For questions 1, 2 and 3 where options were referral with or without treatment of a low-risk woman misclassified as high-risk (false positive) vs. no- or minimal- action such as continuing routine community based antenatal care of a high-risk

woman misclassified as low-risk (false negative), the preference was for the false positive situation 78% (range: 60-100%) of the time. The remaining questions assessed the impact of addition of specific clinical attributes such as preterm gestational age or severe systolic blood pressure on the decision of appropriate community treatment, but due to low response rate (only five respondents completed this section of the questionnaire), interpretation of the results is difficult. For all of questions 4-7, responses (where attributes of gestational age and and systolic blood pressure were varied in the choice pairs) were balanced in their choice between the two options suggesting issues with the wording of the questions themselves or inappropriate use of attributes. Table 4-1 presents a summary of all responses and missing values.

After completion of the questionnaire, results were summarized and presented to the working group with a miniPIERS risk stratification table that was based on interim study data (a version of Table 3-6). Consensus within the miniPIERS working group as a whole was reached; setting the threshold of the miniPIERS predicted probability at ≥25% as a trigger for initial treatment and transfer to a higher level care facility. This decision point was felt to provide adequate sensitivity (approximately 50%) without an overburden to the health system due to greater than 10% false positive referrals.

4.3.2 Design of the POM decision algorithm

Table 4-2 presents a summary of the group activity performed to define additional triggers for treatment and referral. The groups described similar thresholds and a consistent set of

measures required for community based assessment of women with HDP. Based on this exercise, the set of parameters presented in Table 4-3 were determined as the minimum set of clinical variables to be included in the first draft of the POM application. These measures included basic demographics, symptoms and signs along with severe emergency obstetric indicators such as loss of consciousness or evidence of stroke. Areas where the groups differed were around inclusion of proteinuria as a trigger for referral or treatment and the range of blood pressure necessitating treatment with one group suggesting any systolic blood pressure greater than 140 mmHg as requiring intervention while another group indicated the threshold should be 170 mmHg. All groups included symptoms and assessment of fetal status as relevant, although none could define the appropriate response to lack of fetal movements. A cut-off for fetal movements as a decision point was later defined as lack of fetal movements for greater than 12 hours based on personal communication with a fetal medicine expert, Jane Norman, who is currently leading a research group at Edinburgh University investigating training of antenatal fetal movement awareness in pregnant women as a preventative tool for stillbirth.

A first draft of a comprehensive decision tree was created as presented in Figure 4-1. This first draft did not include a proteinuria threshold because at this point in the development process, no consensus as to the appropriate threshold had been reached. The first version did include a threshold based on pulse oximetry due to the results of the fullPIERS study[119] but after review this was removed from the decision algorithm as pulse oximeters are rarely available in low-resourced setting facilities, and definitely not in community level clinics.

This first draft of the decision algorithm was presented along with the trigger questionnaire in Appendix D as a tool to initiate iterative discussion at the subsequent PRE-EMPT meeting in November 2012. Responses were documented in meeting minutes alone and were not collected on the questionnaire forms themselves. All working group members who participated in the discussions agreed that severe hypertension (systolic ≥ 160 mmHg), miniPIERS probability ≥25%, signs of eclampsia, stroke or severe vaginal bleeding should trigger an immediate and urgent response from the community health workers. Areas that required further discussion before group consensus could be reached were the appropriate proteinuria cut-off requiring referral or treatment and which treatment to give to women found with vaginal bleeding prior to delivery. After review by the PRE-EMPT working group the first draft of the decision tree was refined and split into a comprehensive decision tree, suitable for use in facilities and including consideration of gestational age for corticosteroid administration, and a simplified decision tree for community based assessment and initial treatment only. Three more iterations of review and feedback were completed over the period of November 2012 – August 2013 to obtain group consensus for the final decision tree models presented in Figure 4-2 and Figure 4-3. The final triggers included in the decision algorithm are summarized in Table 4-4.

4.3.3 Assessment of accuracy of the decision algorithm using the miniPIERS data

When the community level decision algorithm was applied to 2081 women from the miniPIERS cohort, 1113 (53.5%) women met criteria for treatment or referral based on the combined miniPIERS, systolic blood pressure and dipstick proteinuria triggers. Of these 1113 women, 297

(26.7%) had an adverse maternal event at any point after admission to hospital. Individually, severe systolic blood pressure was responsible for the majority of referral indications, occurring in 1060 (50.9%) of the cohort. Sensitivity, specificity and accuracy of the individual decision points used in the decision algorithm along with the complete set of clinical decision points is presented in Table 4-5. When all clinical decision points are considered, the algorithm identifies women who would go on to suffer an adverse maternal outcome with 74.1% [95% CI 69.4% - 78.2%] sensitivity and 51.4% [95% CI 49.0% - 53.8%] specificity and an overall accuracy of 55.8%. Using miniPIERS alone as a decision point resulted in increased accuracy (81.8%) due to an improvement in specificity (95.2% [95%CI 94.0% – 96.1%]) but was associated with a significant decrease in sensitivity (25.9% [95%CI 21.8% - 30.6%]).

4.4 Discussion

4.4.1 Main findings

In this study, we created a decision algorithm for community-based assessment of women with hypertensive disorders of pregnancy in low-resourced settings based on expert clinical opinion, and validated using data from the miniPIERS cohort. This algorithm includes decision points for referral and treatment based on the miniPIERS model (≥25%), systolic blood pressure (≥160 mmHg), dipstick proteinuria (≥4+), presence or absence of fetal movements and severe emergency obstetric events such as significant vaginal bleeding or eclampsia. Treatment recommendations generated based on these decision points were defined by the WHO recommendations for care of women with pre-eclampsia or eclampsia [7]. Group consensus

around use of these decision points to guide care was established including a convenience sample of relevant stakeholders such as clinicians and researchers from the settings for which this tool is designed. It was clear from the group exercises completed that stakeholder's preference were for a community assessment tool with high sensitivity at the expense of some false-positive referrals.

The accuracy of the tool for identifying a high-risk population defined using the miniPIERS cohort was investigated and found to be moderate when all decision points were included (55.8%) but resulted in the highest sensitivity (74.1% [95% CI 69.4 – 78.2]). Using the miniPIERS decision rule alone resulted in the greatest accuracy (81.8%) but had low sensitivity (25.9% [95% CI 21.8 – 30.6]), meaning many high-risk women would be missed if we relied on the miniPIERS decision rule alone when screening a population of women with hypertensive disorders of pregnancy for risk of adverse maternal outcomes. As highlighted by the DCE, the primary aim of this decision aid is to correctly identify the high-risk cases so that they may receive timely and life-saving interventions. Achieving a high sensitivity using the complete decision algorithm is necessary to meet this target.

4.4.2 Strengths and limitations

One of the strengths of this study is the utilization of a needs-based approach to design a tool for use in LMICs [54,66]. This approach involved inclusion of relevant stakeholders from the target implementation setting in the design and development of the decision algorithm. By including these stakeholders throughout the entire application design process, the resulting tool will be

more easily acceptable for use in the community settings within our target countries. Clinicians will trust the recommendations provided to the community health workers and be able to support the use of this tool in their community and the required task-shifting of clinical evaluation of women with HDP to community-based health workers. Although use of these stakeholders in the design process is a great strength of the study, it is important to note that the group of stakeholders involved was recruited based on convenience and due to their involvement in the study working group. The opinions expressed by them may not be representative and generalizable to all LMIC settings.

A second strength of the study is the use of evidence-based definitions of risk on which to make treatment and referral decisions. The miniPIERS model represents an improvement on previous definitions of HDP disease severity in that it is a fully-validated model that allows calculation of risk for individual women based on the probability of an adverse maternal outcome occurring within 48 hours. Creating a decision rule around this risk estimate limits the individualized application of the model by creating dichotomous risk groups but has the benefit of reducing the complexity of interpretation of risk and allows the model to be easily used by our target community-health worker user group as a guideline for further intervention and management of women in a community setting. We are still reliant on expert opinion for several decision points, making the final tool a mix of evidence-based and opinion-based guidelines but it is an improvement to previous guidelines which used purely opinion-based definitions of risk and disease severity and were often defined by experts from high-resourced settings alone (for example [3,4,82,120]). In addition, we were able to build on the evidence for use of the additional

decision points by completing an initial assessment of their accuracy and prognostic performance using the miniPIERS cohort.

A limitation of this study is the use of the miniPIERS cohort to estimate the accuracy of the decision points defined. Although this cohort allowed us to estimate the impact of this tool on expected referral rates from a community setting to tertiary care centres in an LMIC setting, the miniPIERS cohort is not representative of all community-based pregnant women. The miniPIERS cohort was restricted to women who were selected based on diagnosis of an HDP and admission to hospital and therefore, the rates of occurrence of decision points included in the decision algorithm would be artificially increased due to this selection bias. When applied to a general obstetric population at the community level, referral and treatment rates would likely be lower because the overall prevalence of signs such as severe hypertension or a high miniPIERS probability will be lower. The selection bias introduced through the use of the miniPIERS cohort makes it impossible to draw conclusion as to the true effect of implementation of this tool on population-level referral and treatment rates.

4.4.3 Interpretation and conclusion

Previous guidelines for the management of women with hypertensive disorders of pregnancy in low-resourced settings, such as those published by the WHO [7], have focused on facility based care and interventions once a diagnosis of pre-eclampsia or eclampsia is made. Although local guidelines for assessment of women with HDP do exist in many of the participating countries'

facilities, this study represents the first time an attempt has been made to define evidence-based guidelines for assessment and management of women with HDP at a community level.

The primary reason that initial management and treatment of women with HDP has previously been confined to facilities is the complex nature of the decisions and assessments recommended for management of these cases as described by existing guidelines such as the one published by the National Institute for Health and Clinical Excellence [2]. Although a recent review of physician response to utilization of community-based health workers for pregnancy care showed a shift to acceptability of this task-shifting process, it was clear that many physicians still had reservations and thought task-shifting should be performed only as needed and when evidence for effectiveness was present [62]. By creating the miniPIERS model, using data from a population of women with HDP in low-resourced settings and including only those assessments and clinical measures that could be easily collected by a community-based health worker, we set the foundation for supporting task-shifting of care of these women to available health workers based on evidence of strong prognostic performance. Creating the POM application in collaboration with health workers in LMICs takes this process one step further by providing a simple and evidence-based tool that can be used by community-based health workers and meets the unique needs of health workers in low-resourced settings.

Based on our initial assessment of the accuracy and prognostic performance of these guidelines, it appears that application of this tool into a community-based antenatal care programme would be associated with a false positive rate of approximately 50%. This is of

concern when the consequence of a false-positive is unnecessary utilization of an already overburdened and under-resourced health system. Although our work suggested that clinicians within these settings appear to prefer increased false-positive referrals in order to correctly identify a higher number of truly high-risk women, the potential impact this high rate of referral could have on the health system as a whole must be considered. Any potential impact and improvement in maternal outcomes at a population level may be removed if the health system in which we are implementing the tool is not able to cope with the additional burden of cases referred or is unable to provide quality care at the facility level. The true impact of this tool in a clinical setting requires further evaluation in an implementation study and decision points defined may require adjustment once the tool is actually in use.

The PIERS on the Move tool is the first such tool to be created specifically for use in a community level antenatal care programme. The final stages of tool development require user-interface design and review by the target users, community health workers. This work has been completed through the PIERS on the Move study in South Africa and is being reported elsewhere (J Lim, submitted September 2014). During the usability studies, 37 nurses and midwives in South Africa (15: Tygerberg Hospital; 22: Frère Maternity) evaluated the user interface between November 2011 – January 2013. Each evaluation involved a user completing a simulated patient evaluation during which they would be observed and any errors or difficulties in use of the tool recorded. During the first round of usability testing, major issues in the functionality of the touch-screen keyboard and date scroll wheels were identified; during the second, major improvements in navigation of the application were suggested; and finally

during the third round, the feedback was satisfactory and only minor improvements to navigation were required. Overall, users felt the application was pleasant and would improve their ability to care for hypertensive women. A final version of the application was developed and is now undergoing pilot testing in Africa and South Asia.

Implementation to test the impact of use of this tool in real clinical situations is now ongoing as part of the CLIP (Community Level Interventions for Pre-eclampsia) cluster randomized controlled trial in Nigeria, Mozambique, Pakistan and India (clinicaltrials.gov ID# NCT01911494)[144].

Table 4-1: Summary of discrete choice exercise responses presenting rate of response and proportion of respondents choosing the false positive option of the choice pair.

Question	Response rate	Proportion choosing false positive scenario n (%)
1	15/15	12 (80%)
2	15/15	9 (60%)
3	15/15	15 (100%)
4	5/15	3 (60%)*
5	4/15	2 (50%)*
6	5/15	3 (60%)*
7	5/15	3 (60%)*

* reflects proportion who responded and chose option a

Table 4-2: Summary of responses from country groups to decision tree development exercise highlighting reasons for antihypertensive use, $MgSO_4$, referral (urgent or non-urgent transport to hospital) and other important measures indicated.

	Reason for antihypertensive use	**Reason for $MgSO_4$**	**Urgent transport**	**Non-urgent transport**	**Additional comments**
1	BP>150/100	BP>=170/120 with proteinuria; hypertension and proteinuria with any symptoms	Seizures; abruption; lack of fetal movement or heartbeat; unresponsive; severe oedema	Preterm with mild disease; hypertension with one or two symptoms; non-severe hypertension with proteinuria	Assessment should include medical history and gestational age

	Reason for antihypertensive use	Reason for $MgSO_4$	Urgent transport	Non-urgent transport	Additional comments
2	BP>160/110; eclampsia or imminent eclampsia; reflexes/clonus	Eclampsia or imminent eclampsia	BP>160/110; eclampsia or imminent eclampsia; reflexes/clonus	BP>140/90 with or without proteinuria	Questions to include in assessment: medication history; fetus number; social history (financial situation and decision maker in family); symptoms; fetal movements; bleeding; allergies/asthma

	Reason for antihypertensive use	Reason for $MgSO_4$	Urgent transport	Non-urgent transport	Additional comments
3	BP ≥160/110	Epigastric pain or loss of vision/severe headache; jaundice; haematuria; BP≥160/110 with protein >1+	Epigastric pain or loss of vision/severe headache; jaundice; haematuria; BP≥160/110 with protein >1+; seizures; PV bleeding; BP≥160/110; uterine tenderness; significant proteinuria (>3+) with BP>140/90	BP≥140/90 with no other signs or symptoms	Include in assessment medication history; fetal movements; fundal height; urine output; facial oedema

	Reason for antihypertensive use	Reason for $MgSO_4$	Urgent transport	Non-urgent transport	Additional comments
4	BP≥140/100	Severe pre-eclampsia; reflexes up; eclampsia; imminent eclampsia	BP>160/100 and proteinuria; any signs or symptoms of severe pre-eclampsia; eclampsia; altered consciousness	BP≥140/90 but less than 160/110	Reassess in 24hrs if BP 130/90 – 140/100 or proteinuria >1+; additional questions for assessment include fetal movements; gestational age and parity
5	Not specified	Not specified	Term gestation with BP≥150/100 or >34 weeks with hypertension, headache or epigastric pain or BP≥160/110	Gestational age <34 weeks with BP 140/90 – 150/100	Tried to differentiate by GA groups <34, 34-37 or >37 weeks

Table 4-3: Variables to include in the final PIERS on the Move mobile phone application

Variable Type	Name
Demographics	Date of assessment, estimated date of delivery, date of birth, address, surname, unique identifier, parity, current medications, marital status, number of fetuses
Symptoms	Chest pain/dyspnoea, epigastric pain, headache, visual disturbances, abdominal pain, severe nausea or vomiting, vaginal bleeding, facial oedema or peripheral oedema
Signs	Systolic and diastolic blood pressure or mean arterial pressure, dipstick proteinuria
Fetal assessments	Fetal movement count or presence of fetal movements within previous 12-24 hours
Outcomes	Mortality, severe morbidity as in miniPIERS

Table 4-4: Summary of decision points and recommended treatment or referral interventions for the PIERS on the Move mobile application when used in a community setting.

Decision point	Recommendation			
	Anti-hypertensive	$MgSO_4$	Urgent referral (within 4 hrs)	Non-urgent referral (within 24 hrs)
miniPIERS probability ≥25%	✓	✓	✓	
Systolic blood pressure between 140 – 159 mmHg				✓
Systolic blood pressure ≥160 mmHg	✓	✓	✓	
Dipstick proteinuria ≥4+			□	
Significant vaginal bleeding with systolic blood pressure ≥140 mmHg		✓	✓	
Signs of recent stroke or eclamptic seizure	✓	✓	✓	
Unconscious or unresponsive			✓	
No fetal movements felt for ≥12 hrs				✓

Table 4-5: Estimated referral rate from POM decsion algorithm with sensitivity, specificity and accuracy of all and individual clinical triggers evaluated

	Number identified n(%)	Number with outcome n(%)	Sensitivity (95% CI)	Specificity (95% CI)	Accuracy (%)
All decision points	1113 (53.5%)	297 (26.7%)	74.1 (69.4 – 78.2)	51.4 (49.0 – 53.8)	55.8%
miniPIERS probability ≥25%	185 (8.9%)	104 (56.2%)	25.9 (21.8 – 30.6)	95.2 (94.0 – 96.1)	81.8%
Systolic blood pressure ≥ 160 mmHg	1060 (50.9%)	280 (26.4%)	69.8 (65.0 – 74.2)	53.6 (51.2 – 56.0)	56.7%
Dipstick ≥ 4+	210 (10.0%)	62 (29.5%)	15.5 (12.1 – 19.5)	91.2 (89.7 – 92.5)	76.6%

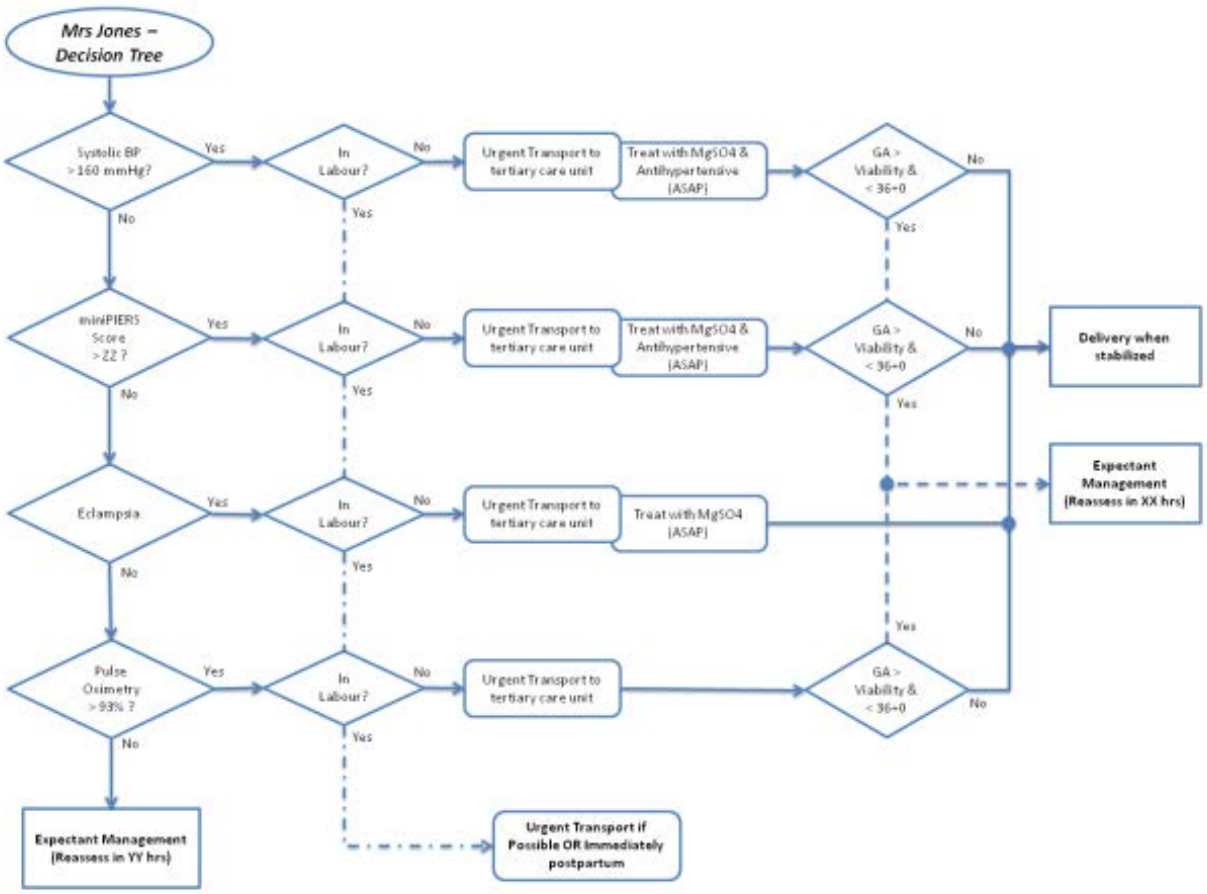


Figure 4-1: First draft decision algorithm for use in POM

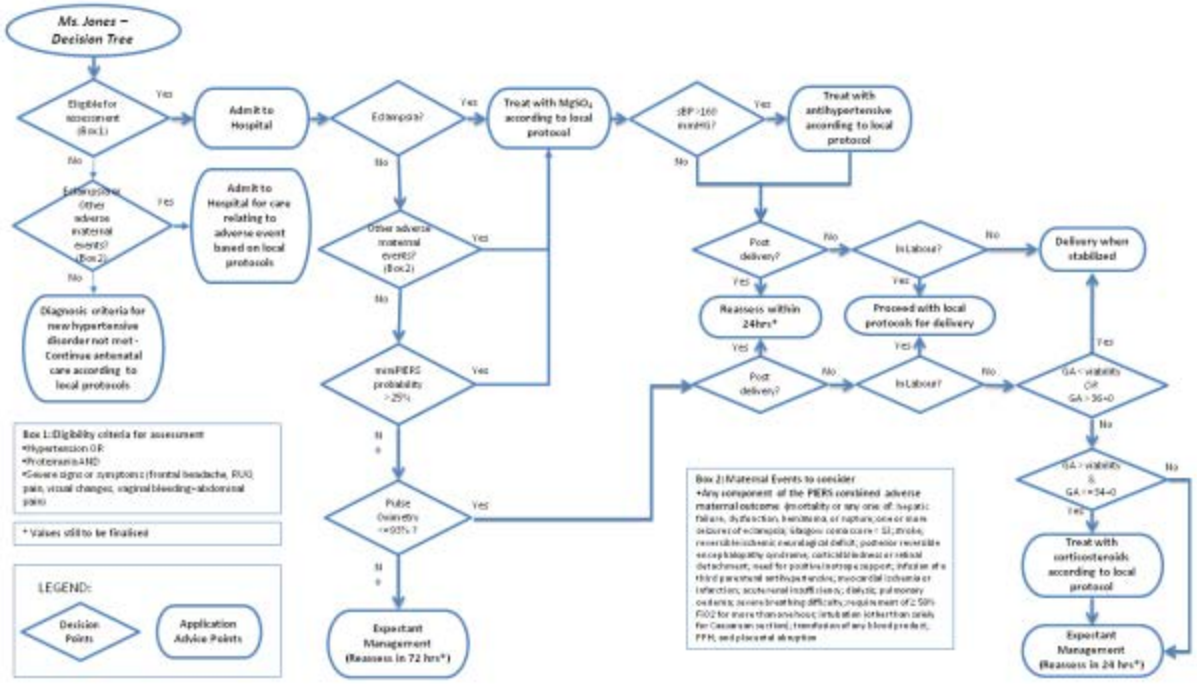


Figure 4-2: Final comprehensive decision tree used in the PIERS in the Move mobile application for use in a facility setting

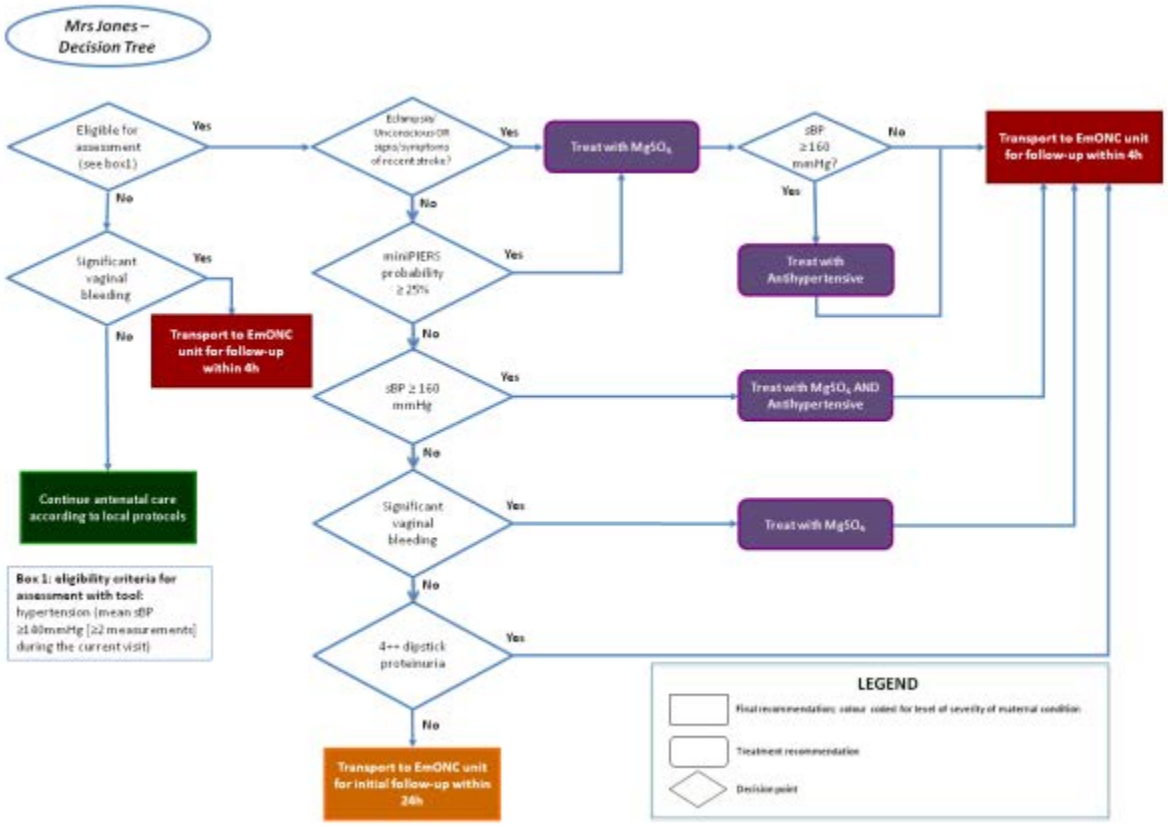


Figure 4-3: Final simplified decision tree for the PIERS on the Move application for use in a community setting

Chapter 5: Assessing the incremental value of blood oxygen saturation measured by pulse oximetry (SpO_2) in the miniPIERS model

5.1 Introduction

In low-resourced settings, current approaches to assess the severity of HDP-related illness and guide clinical decisions are based on assessment of blood pressure and symptoms alone. In the previous Chapter, development of a decision algorithm incorporating WHO treatment recommendations [7] with clinical decision points used to initiate treatment defined by local stakeholders and a decision rule around the use of the miniPIERS model was described.

As presented in Chapter 3 of this dissertation, the goal of the miniPIERS project was to reduce adverse pregnancy outcomes associated with the HDP by providing community-based health workers in low-resourced settings with an evidence-based and low-cost tool to improve risk stratification and management of pre-eclampsia. The miniPIERS model includes symptoms and signs (parity, gestational age at assessment, chest pain/ dyspnoea, headache/ visual disturbances, vaginal bleeding with abdominal pain, systolic blood pressure and dipstick proteinuria) and based on measures of these variables, allows a health workers to determine the risk of adverse pregnancy outcomes occurring within 48 hours of assessment of the hypertensive woman [135]. Although the miniPIERS model as developed and validated shows promise, improvements in the model's accuracy may be possible with the addition of more sensitive risk markers. We have previously shown that blood oxygen saturation (SpO_2) measured by pulse oximetry is a significant independent predictor of risk of complications in

women with pre-eclampsia in an institutional setting [100]. Perturbations in SpO_2 level in the hypertensive women likely reflect the consequences of endothelial dysfunction that is characteristic of maternal hypertensive disorders leading to increased permeability of the pulmonary vasculature and impaired pulmonary diffusion capacity [21]. Given the recent development of a low-cost mobile phone based pulse oximeter, the Phone Oximeter [145], the objective of the study described in this chapter was to assess the incremental value of adding SpO_2 to the miniPIERS model.

5.2 Methods

5.2.1 Study population and design

We conducted a prospective cohort study of women admitted to a participating institution with new (onset after 20 weeks' gestation) or chronic hypertension [blood pressure (BP) ≥140/90mmHg (at least one component, twice, ≥4 and up to 24 hours apart, after 20 weeks)] during pregnancy, with or without proteinuria or other adverse conditions. The participating Institutions were: (i) Tygerberg Hospital, Cape Town, South Africa; (ii) Aga Khan University Hospital and its secondary level hospitals at Garden, Karimabad and Kharadar, and Jinnah Post-graduate Medical College, Karachi, Pakistan; and (iii) Aga Khan Maternity & Child Care Centre, and Liaqat University of Medical Sciences, Hyderabad, Pakistan. Ethics approval for this study was obtained from each participating institution's research ethics board as well as the clinical research ethics board at the University of British Columbia.

Data collected for this study included demographics (parity, gestational age, maternal age, and medical history), symptoms (headache, visual disturbances, chest pain, dyspnoea, abdominal pain with vaginal bleeding, epigastric pain, nausea and vomiting) and clinical signs (blood pressure, diptick proteinuria and SpO_2) as shown in the data collection form in Appendix A. At Tygerberg Hospital, Cape Town, South Africa, data were collected using a version of the PIERS on the Move (POM) mobile health (mHealth) application similar to that described in Chapter 4 but with the addition of the Phone Oximeter [145](merged). The POM application was designed in collaboration with nurses and midwives in South Africa, Pakistan, India and Nigeria [142] as a decision aid for community health workers incorporating the miniPIERS model and other clinical indicators, a novel mobile phone based pulse oximeter [145] and the WHO recommendations for management of women with pre-eclampsia and eclampsia [7]. A detailed description of development of the decision algorithm used in the POM tool was presented as Chapter 4 of this dissertation. In this study, POM was used only as a data collection instrument; treatment and management of women was not influenced by the POM application. The study staff assessed consenting women and collected relevant clinical data every four days during their inpatient stay. At the Aga Khan University Hospitals in Karachi and Hyderabad, Pakistan, data were abstracted from medical records of women admitted for care due to a HDP as part of the original miniPIERS study using the data collection form provided in Appendix A.

In both settings, the frequency of evaluations and timing in relation to hospital admission was kept consistent and followed the hospital's mandated standard of care. For the purpose of this study, data from the first clinical assessment after admission to hospital were used. If a relevant

measure was missed during the first assessment, data from subsequent assessments occurring within 24 hours of admission were used to resolve any missing values.

At Tygerberg Hospital where the POM application was used for data collection, additional information on the time taken to complete an evaluation was collected. The research midwife who performed all evaluations with the POM application was also surveyed at the end of the study to understand her impressions of feasibility and acceptability of use of the tool in a clinical setting. The survey tool used for collection of the research midwife's feedback is provided as Appendix F.

5.2.2 Main outcome measures

The primary outcome for this study was a composite adverse maternal outcome, defined as maternal mortality or one or more of serious central nervous system, cardiorespiratory, renal, hepatic, haematological or other morbidity and is the same outcome as previously described in Chapter 3 of this dissertation. A list of components of the adverse maternal outcome with full definitions is provided in Appendix B. The components of the composite outcome were determined by Delphi consensus [125] for the purpose of the original fullPIERS model development and validation project [119]. Data were collected on the occurrence of all outcome components at any time during admission but for the purpose of this study, only those that occurred within 48 hours of admission were considered. All study sites were instructed to collect information on any "other" adverse events the woman experienced during pregnancy or immediately postpartum as part of the regular data collection process. This was done to ensure

balanced reporting of events across all sites. Any reported "other" events were adjudicated by the study Working Group during regular meetings, at which time the decision was made whether or not to include the reported outcome as a study outcome.

5.2.3 The miniPIERS model

The published miniPIERS equation was used to calculate a linear predictor variable for all women in the study cohort. This equation is: miniPIERS linear predictor(lp)= -5.77 + [$-2.98x10^{-1}$ x indicator for multiparity] + [(-1.07)x log gestational age at admission] + [1·34 x log systolic blood pressure] + [($-2·18x10^{-1}$)x indicator for 2+ dipstick proteinuria] + [($4·24x10^{-1}$) x indicator for 3+ dipstick proteinuria] + [($5.12x10^{-1}$) x indicator for 4+ dipstick proteinuria] + [1·18 x indicator for occurrence of vaginal bleeding with abdominal pain] + [($4.22x10^{-1}$) x indicator for headache and/or visual changes] + [$8.47x10^{-1}$ x indicator for chest pain and/or dyspnoea] [135]. The predicted probability of adverse maternal outcome was determined using the following equation: $p=e^{lp}/1+e^{lp}$

A threshold of 25% predicted probability was used to define the high-risk population based on the optimal threshold identified during development and validation of the miniPIERS model as described in Chapter 3 and presented in the published version of the miniPIERS study [135].

5.2.4 Sample size requirements

Simulation studies have demonstrated that the sample size requirement to identify issues with model calibration (such as over or underestimation of risk) is 100 cases with an adverse

outcome and 100 cases with no adverse outcome [112,126]. Therefore, data collection was planned to continue until a minimum of 100 adverse outcomes had occurred within the study cohort.

5.2.5 Statistical analysis

Demographics for women from each study setting were described using means and standard deviations or medians with interquartile ranges, when not normally distributed, for continuous variables and based on counts with frequencies for categorical variables.

The association between SpO_2 and the composite adverse maternal outcome was assessed using logistic regression. Multivariable logistic regression was used to further adjust for the other predictor variables in the miniPIERS model, as these are known to be significantly associated with risk of adverse maternal outcome. The ability of SpO_2 to discriminate between women who did and did not meet the outcome criteria was assessed based on the area under the curve (AUC) of the receiver operating characteristic (ROC) curve [129]. In order to confirm that the observed relationship was generalizable to non-respiratory outcomes, the relationship between SpO_2 and the adverse maternal outcome was assessed against both the complete composite adverse maternal outcome and a restricted adverse outcome where cardiorespiratory events were removed from the composite outcome. A final sensitivity analysis was performed to assess the effect of SpO_2 on outcome in each study site to rule out any possibility of confounding by centre.

Recalibration and extension of the miniPIERS model to include SpO_2 was performed by fitting a new model using the study cohort that included two variables: (1) the linear predictor from the

original miniPIERS model; and (2) a continuous measure of SpO_2 [111]. This simple method of updating the model was chosen to make the best use of the previously validated miniPIERS model. Should the pulse oximeter fail to work in the field, the model could still be used by reverting back to the original miniPIERS equation as the parameters remain fixed within this recalibrated and extended model.

Performance of the extended model and the original miniPIERS model in this cohort were assessed for discrimination ability based on the AUC ROC and calibration using the Hosmer-Lemeshow goodness of fit test. The two models were then compared based on stratification capacity and classification accuracy using a reclassification table [128,146,147]. Net change in model performance based on inclusion of SpO_2 was also assessed using a net reclassification index (NRI) and by evaluating the change in true- and false- positive rates [146,148-150] at the previously published 25% predicted probability threshold for a positive test. The NRI is calculated as the improvement in classification for each of the sub-groups of the study population with and without events using the formula: NRI=((P(up|event)-P(down|event))+((P(down|nonevent)-P(up|nonevent)); where 'up' refers to reclassification by the extended model to the higher-risk group and 'down' refers to reclassification by the extended model to the lower-risk group. The sensitivity, specificity, positive and negative predictive values and positive and negative likelihood ratios of various cut-points of predictive probability were also calculated for both models to compare performance across multiple risk categories.

For assessment of model performance, an AUC ROC of >0.70 was considered 'good' according to established standards in interpretation [129]. The following categories for interpretation of the likelihood ratios were used: informative (LR<0·1 or >10); moderately informative (LR 0·1-0·2 or 5-10); and uninformative (LR 0·2-5) [130]. As described in Chapter 3, we had previously determined that sensitivity was considered clinically more relevant as a marker of performance of the miniPIERS model as a screening tool. For that reason, improvement in model sensitivity when adding the new marker was considered of primary importance in this study.

All statistical analyses were performed using STATA v11·0 (StataCorp, College Station, TX, USA).

5.3 Results

Between 1 January 2011 and 31 March 2012, 617 women were recruited to the study in Pakistan, while 235 women were recruited in South Africa between 1 November 2012 and 31 December 2013. These two groups combined to create a study cohort of 852 women of whom 119 (14.0%) experienced one or more component of the composite adverse maternal outcome within 48 hours of admission. A total of 147(17.3%) women experienced one or more component of the composite adverse maternal outcome at any time during hospital admission. The women recruited from South Africa tended to be earlier in gestation and more often admitted with a diagnosis of pre-eclampsia rather than gestational hypertension alone. This more severe case-mix is reflected in the increased use of corticosteroids and magnesium sulphate ($MgSO_4$) in the Tygerberg cohort. The overall rates of both maternal and perinatal adverse outcomes were comparable between sites, as presented in Table 5-1.

The most common adverse outcomes in the cohort were need for blood transfusion, pulmonary oedema and postpartum haemorrhage. There were no maternal deaths in the study population but there were 14 cases of eclampsia. A complete description of outcomes occurring in the cohort at any time and within 48 hours of admission is provided in Table 5-2.

Increased SpO_2 was significantly associated with a decreased risk of adverse maternal outcome OR 0.65 [95% CI 0.59, 0.72]. This association remained after adjustment for all other miniPIERS predictor variables of parity, gestational age at admission, chest pain/ dyspnoea, headache/ visual disturbances and abdominal pain with vaginal bleeding, systolic blood pressure and dipstick proteinuria (univariate (adjusted OR 0.70 [95% 0.62, 0.79]). As shown in Table 5-3, women with $SpO_2 \leq 93\%$ were 30.7-fold [95% CI 13.9, 67.7] more likely to have an adverse outcome than women with $SpO_2 > 97\%$. The results were also similar when assessing the effect of SpO_2 on outcome in the data from Pakistan and South Africa individually with ORs of 0.68 [95% CI 0.61, 0.76] and 0.53 [95% CI 0.42, 0.68], respectively.

SpO_2 alone, and when adjusted for the other miniPIERS predictor variables, resulted in an AUC ROC of 0.73 [95% CI 0.68, 0.78] and 0.81 [95% CI 0.76, 0.86], respectively. When a sensitivity analysis was performed using only non-cardiorespiratory outcomes, SpO_2 maintained its discriminatory ability with an AUC ROC of 0.69 [95% CI 0.64, 0.74] when unadjusted and AUC ROC of 0.75 [95% CI 0.69, 0.81] when adjusted for the other miniPIERS risk factors.

When the original miniPIERS model was applied to our study cohort, the AUC ROC was 0.78 [95% CI 0.73, 0.83] and Hosmer-Lemeshow goodness of fit p-value was 0.16. After model

recalibration and extension, the AUC ROC was 0.80 [95% CI 0.75, 0.85] with a goodness of fit p-value equal to 0.70. The ROC curves for both models are presented as Figure 5-1.

Comparison of the models based on their ability to classify women correctly as high-risk using a threshold of predicted probability of 25% is presented in Table 5-4. When extending the model to include SpO_2, the number of women who are correctly reclassified into the high-risk group who did in fact have an adverse maternal outcome is 22 (18.5% of all 119 cases with an adverse outcome), there are also two women who were incorrectly reclassified by the extended model as low-risk who did in fact suffer an adverse outcome. This results in an overall change in true-positive rate of 0.17. In the low-risk group presented in Table 5-4, there were four women correctly reclassified as low-risk with the extended model and 38 women incorrectly reclassified as high-risk who did not suffer an adverse maternal outcome resulting in an overall change in true-negative rate of -0.05. The overall rates of change in true-positive and true-negative rates are combined to calculate an NRI of 0.12.

Sensitivity, specificity, positive and negative predictive values and likelihood ratios for cut-points of 20%, 25% and 30% predicted probabilities are presented for comparison in Table 5-5. The extended model had a trend towards improved sensitivity and NPV at all risk thresholds evaluated but decreased specificity, PPV and likelihood ratios. For example, at the 25% predicted probability threshold comparing the original model to the extended model resulted in sensitivity of 32.8% vs 49.6%; specificity of 96.2% vs. 91.5%; PPV of 58.2% vs. 48.8%; NPV of

89.9% vs .91.8%; positive likelihood ratio of 8.6 vs. 5.9; and negative likelihood ratio of 0.7 vs. 0.6, respectively. For all measures, confidence intervals overlapped.

When surveyed to determine feasibility of use of the POM application in a clinical setting the research midwife reported that the application was "*easy to use and learn*"; the only complication she experienced while using the application was that in some cases it was difficult to achieve the necessary Signal Quality (SQI) when measuring the SpO_2. She also reported: *"Patients appreciated the personal attention and a time to ask questions regarding their condition. In this sense the apparatus was still linked to the care of the midwife. It was not our impression that the application was appreciated for its' own merits by our patients."* After review of more than 500 evaluations completed on the 235 women evaluated over the course of the study, the average time to complete an evaluation was found to be 5 minutes and 57 seconds [sd 6 minutes and 17 seconds].

5.4 Discussion

5.4.1 Main findings

In this study we found that SpO_2 was a significant predictor of risk of adverse maternal outcomes in women with a HDP, with a threshold of ≤ 93% SpO_2 associated with a 30-fold increase in risk compared with women with normal oxygen saturation, at or above 98%. This association was consistent across study settings and after adjustment for other risk factors included in the miniPIERS model. A sensitivity analysis performed to assess the effect of SpO_2 as a predictor of non-respiratory outcomes alone also demonstrated a consistent predictive effect.

This suggests reduced pulmonary function measured as decreased SpO_2 is a marker of severe disease in general and is not specific to risk for pulmonary oedema alone.

Recalibration of the miniPIERS model in the study cohort improved model calibration based on the Hosmer-Lemeshow goodness of fit test and maintenance of discriminatory ability as shown by the AUC ROC of 0.781 [95% CI 0.729, 0.832]. Extension of the model to include SpO_2 produced very similar results: an AUC ROC of 0.798 [95% CI 0.752, 0.846]. The inclusion of SpO_2 increased the model's ability to identify true-positive cases (increasing from sensitivity of 32.8% to 49.6%) and had a positive net reclassification index value of 0.122, but at the expense of an increase in the number of false-positive cases (with specificity decreasing at all risk thresholds examined (e.g. from 96.2% to 91.5%)).

5.4.2 Strengths and limitations

Strengths of this study include the large sample size and high quality of the data collected. In a cohort of over 800 women, we were able to collect complete data for all cases enrolled.

Another major strength of this study is in the methodology used. Rather than simply presenting an updated model that includes SpO_2, we have carefully considered the incremental value of the new predictor. This is particularly important given the target low-resourced setting in which we want to implement the miniPIERS tool. In this case, we demonstrate that the improvement found would be an approximately 20% increase in a health worker's ability to detect high-risk patients (true-positives), but at the expense of an increase in the number of low-risk women incorrectly classified as high risk (false positives). Addition of SpO_2 would only be warranted in

these settings if the added resource requirements, in the form of an accompanying increase in false positive cases, could be properly balanced with clinically relevant improvements in outcomes for the additional high-risk patients identified. It will depend on the local setting's resource availability as to whether this is manageable.

The use of data from two different sites, collected in two distinct ways (retrospective chart review vs. prospective and direct data capture) is a major limitation of this study. Combining the populations from our two study sites was necessary in order to meet sample size requirements for model recalibration but may have introduced additional biases due to differential misclassification of predictors or measurement errors that may have occurred when using the clinically available pulse oximeter and data from the medical record in Pakistan. We attempted to reduce these biases by ensuring data collection in South Africa was at the same time-points as that in Pakistan using the same definitions of parameters. We also ensured that clinical practice guidelines for treatment and management of women with hypertensive disorders of pregnancy were consistent across both settings. The consistent effect of the predictors within each cohort supports combining data from the two sites. In addition, using data from two sites may also be considered a strength of the study as it increases the generalizability of our results to multiple locations.

A second limitation is in the available methodology for comparing the original and extended model. Currently, the determination of net reclassification index and use of reclassification tables are the recommended methodologies for assessing model improvement when adding a

new risk factor to a diagnostic or prognostic model, but interpretation of the NRI is still an area of investigation and debate [151]. This study demonstrates the NRI's main limitation as we have shown that the improvement in the model comes entirely from an increase in true-positive rate. If we were to interpret the NRI alone we may have missed the fact that there is actually a decrease in model true negative rate with the addition of the new predictor. Similarly, it is not possible to compare models based on the AUC ROC particularly due to the relatively greater importance of sensitivity over specificity for this tool [152,153]. Therefore, interpretation of these results is dependent on weighing the clinical and resource implications of an increase in true-positive vs. true-negative rate.

An indirect limitation of this research is the limitation of the study primary outcome to maternal events alone. In this study we have focused solely on maternal outcomes as our primary measure of effect as this is the focus of the original miniPIERS model. A future direction that should also be considered is assessment of the effect of pulse oximetry as a measure of neonatal prognosis. The impact of introduction of a tool such as the phone oximeter would be greatly strengthened if it were found to be a significant predictor of stillbirth or neonatal survival; unfortunately, the miniPIERS study was designed with this endpoint in mind.

A final limitation is that due to the health workforce shortage at Tygerberg Hospital in South Africa, we were limited to use of a single research midwife for all data collection for this site. Because of this, the generalizability of our assessment of the clinical feasibility and acceptability of the POM application is limited. It may be that the time taken to complete evaluations by

other health workers would be longer or shorter, depending on their experience with mobile phones and training in antenatal care. In future assessments of the tool, inter-rater variability in use should be estimated but at this early stage in development having one consistent user is sufficient to demonstrate potential usability.

5.4.3 Interpretation

The performance of the miniPIERS model with or without SpO_2 found in this study is consistent with that found during development and external validation of the original model where the AUC ROC was 0.768 [95% CI 0.735-0.801] in the development dataset and 0.713 [95% CI 0.658-0.768] on external validation [135]. The independent effect of SpO_2 on maternal outcome is also consistent with previously published results where SpO_2 ≤93% was associated with an approximately 18-fold increase in risk of the same combined adverse maternal outcome (95% CI 8.1-40.1) in a population of women with pre-eclampsia admitted to tertiary perinatal units in high-resourced settings [100]. This supports the conclusion that the effects seen in this study are accurate and generalizable to other settings.

The main clinical consideration in the interpretation of this study's result is the trade-off required between true and false positives when applying a tool such as miniPIERS as a screening test for individual risk within a population. If the risk threshold used to define the high-risk group is set too high, the consequence would be that many women who truly need referral and further treatment would be missed; if the risk threshold is set too low, there will be

an increase in false positives which may result in overburdening the higher-level health facilities with women who do not require more care.

During the development of the original miniPIERS model we undertook a survey of clinical consultants at all study sites and within the study working group to determine the priority (sensitivity vs. specificity) when setting the risk threshold used to define the high-risk population, as described in the previous chapter. We focus on the use of the 25% predicted probability as the optimal risk threshold as it was felt to demonstrate adequate performance as a rule-in test without increasing the false positive rate above approximately 10%. This process is described in detail in the previous chapter. A similar pattern in results in terms of sensitivity and false positive rate was found in this study.

5.5 Conclusion

The miniPIERS model, when applied as a screening tool in a pregnant population, should allow health workers to accurately stratify women into useful risk groups. Addition of SpO_2 may improve the clinical utility of this stratification process by improving sensitivity of the model. By stratifying the population based on evidence-based risk using the miniPIERS model, health workers can assess individual women for risk of complications related to hypertension in pregnancy occurring within 48 hours of assessment. This timeframe would allow decisions to be made regarding treatment and referral to a higher level of care that could make the difference in a woman's life. This is currently the only tool of its kind available for this purpose. Our pilot

assessment of the POM tool in a clinical setting in South Africa suggests it would be feasible to use by mid-level health workers in LMICs.

Addition of SpO_2 to the model does confer a net improvement in model accuracy based on an increase in the model's sensitivity. This increase in sensitivity is accompanied by an increase in false positive rate that would result in additional health resource use. Before this tool can be scaled-up, further research is required to assess the true impact on pregnancy outcomes of introduction of this tool to an antenatal care program in a low-resourced setting. This implementation study is now underway as part of the PRE-EMPT (Pre-eclampsia and Eclampsia, Monitoring, Prevention and Treatment) initiative, called the Community Level Interventions for Pre-eclampsia (CLIP) study [144,154].

Table 5-1: Demographics and clinical status at admission for women admitted to the study from either Pakistan or South Africa.

Characteristic	Pakistan cohort (N= 617 women)	South African cohort (N= 235 women)
Demographics		
Maternal age at EDD (years)	29 [26, 33]	27 [23, 33]
Gestational age at eligibility (weeks)	37.2 [35.4, 38.2]	34.6 [30.0, 37.9]
Multiple pregnancy	13 (2.1%)	1 (0.4%)
Parity ≥ 1	320 (51.9%)	126 (53.6%)
HDP description		
Pre-eclampsia†	343 (55.6%)	173 (73.6%)
Other HDP	274 (44.4%)	62 (26.4%)
Clinical Measures within 24 hours of admission		
Systolic BP	150 [140, 160]	146 [140, 160]
Diastolic BP	100 [90, 110]	96 [90, 101]
Dipstick proteinuria	2+ [trace, 2+]	2+ [1+, 3+]

Characteristic	Pakistan cohort (N= 617 women)	South African cohort (N= 235 women)
SpO_2	97 [95, 98]	97 [96, 98]
Interventions		
Corticosteroid administration	146 (23.7%)	143 (60.9%)
Antihypertensive medications administered	596 (96.6%)	234 (99.6%)
$MgSO_4$ administered	231 (37.4%)	186 (79.1%)
Pregnancy outcomes		
Intrauterine fetal death (≥20+0 wk and/or ≥500g)	59 (9.6%)	21 (8.9%)
Neonatal death (before discharge)	22 (3.4%)	7 (3.0%)
Maternal adverse outcome (within 48 hours of admission)	91 (14.7%)	28 (11.9%)

†pre-eclampsia defined as hypertension (blood pressure greater than 140/90 mmHg with proteinuria greater than 2+ on a dipstick test

Table 5-2: Maternal adverse outcomes occurring in the study cohort. Full definitions of all outcomes are available as Appendix A.

One or more of maternal morbidity or mortality:	within 48h	any time
TOTAL n(%)	**119 (14.0%)**	**147(17.3%)**
Maternal death	0	0
Central nervous system		
Eclampsia (≥1)	10	14
Glasgow coma score <13	7	8
Stroke or reversible ischaemic neurological deficit	1	1
Cortical blindness or retinal detachment	3	3
Posterior reversible encephalopathy	0	1
Cardiorespiratory		
Positive inotropic support	1	1
Infusion of a 3rd parenteral antihypertensive	0	0
Myocardial ischaemia/infarction	1	3
≥50% FiO2 for >1hr	1	2
Intubation (other than for Caesarean section)	3	3
Pulmonary oedema	23	32
Haematological		
Transfusion of any blood product	46	53
Platelets <50 x 10^9/L with no transfusion	2	2
Hepatic		
Dysfunction	1	1
Haematoma/rupture	0	0
Renal		
Acute renal insufficiency (creatinine > 150µmol/L; no pre-	4	4

One or more of maternal morbidity or mortality:	within 48h	any time
existing renal disease) (creatinine>200µmol/L; pre-existing renal disease)		
Dialysis	0	1
Obstetric outcomes		
Placental abruption	2	7
Postpartum hemorrhage	24	26
Other adverse events		
Severe ascites	15	20
Other*	3	5

*includes two cases of pulmonary embolism, two cardiac arrests and one case of ruptured uterus

Table 5-3: Adverse outcome rate by strata of SpO2 (n=852) and odds ratio for occurrence of adverse maternal outcome in each strata compared to odds of outcome in women with spO2 greater than 97%

SpO_2 level	Women in the cohort [n(%)]	Women with adverse outcomes [n(%*)]	Univariate OR [95% CI]
≤ 93%	38 (4.5%)	25 (65.8%)	30.7 [13.9, 67.7]
94-95%	153 (18.0%)	34 (22.2%)	4.6 [2.6, 8.0]
96-97%	271 (31.8%)	37 (13.7%)	2.5 [1.5, 4.4]
≥98%	390 (45.8%)	23 (5.9%)	reference

*denominator used is number of women in cohort with SpO_2 in the given range

Table 5-4: Classification table comparing miniPIERS model with and without addition of SpO2

Model without SpO_2	**Model with SpO_2**		**total**
Women with events	**0-24.9%**	**≥25.0%**	
0-24.9%	58	22	80
≥25.0%	2	37	39
Total	60	59	119
Women without events			
0-24.9%	667	38	705
≥25.0%	4	24	28
Total	671	62	733

Table 5-5: Performance measures for the original miniPIERS model and the extended model at various cut-points of predicted probability used to define a positive test.

	25%		15%		35%	
	Original model	**Extended model**	**Original model**	**Extended model**	**Original model**	**Extended model**
Sensitivity (95% CI)	32.8 (24.6-42.1)	49.6 (40.3-58.8)	50.4 (41.2-59.7)	68.1 (58.8-76.1)	20.2 (13.6-28.7)	39.5 (30.8-48.9)
Specificity (95% CI)	96.2 (64.5-97.4)	91.5 (89.2-93.4)	91.7 (89.4-93.6)	77.9 (74.7-80.8)	98.0 (96.6-98.8)	96.3 (94.6-97.5)
PPV (95% CI)	58.2 (45.5-69.9)	48.8 (39.6-58.0)	49.6 (40.4-58.8)	33.3 (27.5-39.7)	61.5 (44.7-76.2)	63.5 (51.5-74.2)
NPV (95% CI)	89.9 (87.4-91.8)	91.8 (89.5-93.6)	91.9 (89.7-93.8)	93.8 (91.5-95.5)	88.3 (85.9-90.4)	90.7 (88.4-92.6)
LR+ (95% CI)	8.6 (5.5-13.4)	5.9 (4.3-7.9)	6.1 (4.5-8.2)	3.1 (2.6-3.7)	9.9 (5.3-18.2)	10.7 (7.0-16.5)
LR- (95% CI)	0.7 (0.6-0.8)	0.6 (0.5-0.7)	0.5 (0.5-0.6)	0.4 (0.4-0.6)	0.8 (0.7-0.9)	0.6 (0.5-0.7)

PPV=positive predictive value; NPV= negative predictive value; LR+= positive likelihood ratio; LR-=negative likelihood ratio

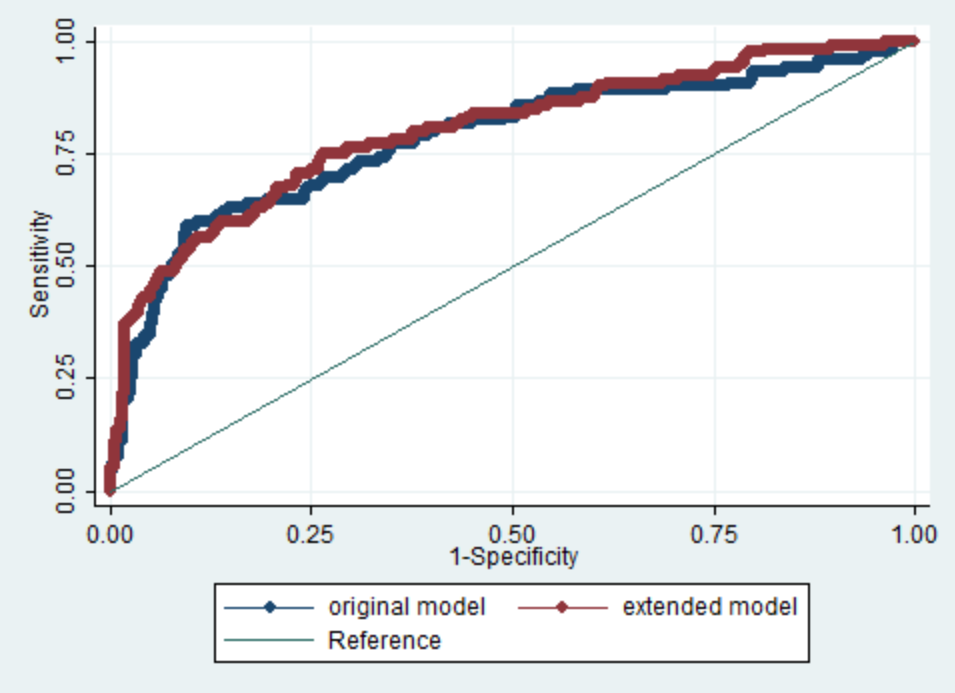


Figure 5-1: Receiver operating characteristic curve for the original (blue) and extended (red) miniPIERS model. Area under the curve for the original model was 0.781 [95% CI 0.729 - 0.832] and for the extended model was 0.798 [95% CI 0.752 - 0.846].

Chapter 6: Assessing overfitting in maternal clinical prediction models: the contributions of variable coding, variable selection and parameter estimation

6.1 Introduction

Clinical prediction models have become widely used in medical practice as tools to improve clinician's ability to either diagnose a disease or determine disease prognosis. Methods for development of clinical prediction models have been described and follow standard procedures including: (1) variable inspection and coding (checking the model assumptions of additivity and linearity of effects and defining required transformations or interaction terms to be tested); (2) selection of predictors (using a fully specified model based on previous literature or a reduced model based on variable selection techniques such as forward or backward elimination); and (3) parameter estimation (using regression) [107,111,117].

A major concern in the development of a clinical prediction model is the occurrence of overfitting, whereby the model's risk prediction equation reflects idiosyncrasies of the study dataset rather than true generalizable relationships. Overfitting has been demonstrated to occur at both the variable selection stage and at the stage of parameter estimation. Overfitting leads to overly optimistic estimates of the model's predictive ability [73,107,112,115].

At the variable selection stage, overfitting in the development of a clinical prediction model comes from fitting a model with too many degrees of freedom during the modeling process [107,112,115]. A general rule of thumb is that there should be 10 events (cases of adverse outcome) per effective variable tested. For example, in a small dataset of 200 participant including 50

cases of adverse outcomes, only 5 effective variables should be tested during the model development process. A simulation study has suggested that this rule of thumb may even be too liberal, as bias in regression coefficients was still present with 10 cases of adverse outcome per effective variable used [155]. This sample size requirement is not always met in published clinical prediction models [156,157] and therefore, overfitting due to insufficient sample sizes remains a substantial issue.

A second aspect of model development that may lead to overfitting is the practice of selecting variables to include in the model based on the strength of univariate predictor effects seen in the dataset under study [111]. Other established sources of overfitting include model selection methods such as backward elimination, which have been shown to add significant bias to estimates of effect even when what is considered an effective sample size is used [113,158]. Finally, additional overfitting may be introduced when variables are coded (either through categorizations of transformations) based on analysis of the predictor - outcome relationship using the dataset under study, although this has not been demonstrated in the literature to date.

The objective of this study was to determine, using data from two prospective cohorts, the extent to which each of the three stages of model development: variable coding; variable selection and parameter estimation contribute to estimates of prediction model optimism and overfitting.

6.2 Methods

6.2.1 Study design and population

Data for this study were taken from two international, multicentre prospective cohorts: (1) fullPIERS [119], which was completed in high-income country, tertiary level obstetric units in Canada, the UK, Australia and New Zealand; and (2) miniPIERS [135], which was completed in low- and middle- income country obstetric facilities in Pakistan, Uganda, South Africa, Brazil and Fiji and is described in detail in Chapter 3.

The fullPIERS cohort included only women admitted with a diagnosis of pre-eclampsia while miniPIERS included women with pre-eclampsia or an 'other' hypertensive disorder of pregnancy (HDP). For the purpose of both cohorts the following definitions were used: pre-eclampsia, defined as i) blood pressure (BP) ≥140/90mmHg (at least one component, twice, ≥4 and up to 24 hours apart, after 20 weeks) and either proteinuria (of ≥2+ by dipstick, ≥300mg/d by 24 hour collection, or ≥30 g/mol by urinary protein:creatinine ratio) or hyperuricaemia (greater than local upper limit of local non-pregnancy normal range), ii) HELLP (haemolysis, elevated liver enzymes, and low platelets) syndrome even in the absence of hypertension or proteinuria [124], or iii) superimposed pre-eclampsia (clinician- defined rapid increase in requirement for antihypertensives, systolic BP (sBP) ≥170mmHg or diastolic BP (dBP) ≥120mmHg, new proteinuria, or new hyperuricaemia in a woman with chronic hypertension); or an 'other' HDP defined as: i) gestational hypertension ((BP) ≥140/90mmHg (at least one component, twice, ≥4 hours apart, ≥20^{+0} weeks) without significant proteinuria); ii) chronic hypertension (BP

≥140/90mmHg before 20^{+0} weeks' gestation); or iii) partial HELLP (i.e. haemolysis and low platelets OR low platelets and elevated liver enzymes).

The fullPIERS dataset consists of 2023 women, 106 (5.2%) of whom suffered an adverse maternal outcome within 48 hours of admission. Demographics and clinical measures collected within 24 hours of admission for women in this cohort comparing those with and without adverse maternal outcomes have been previously published [122]. A complete breakdown of outcome events to occur in this dataset has been previously published [119]. The miniPIERS cohort included a population of 2081 women, 261(12.5%) suffering one or more component of the adverse maternal outcome. The demographics of women in this dataset have been described in Chapter 3 (Table 3-1). Occurrence of adverse outcome within the cohort is described in Table 3-2).

6.2.2 Data collection procedures and missing data

Data were collected using consistent protocols and data collection forms as provided in Appendix A of this dissertation. Candidate predictor variables included demographics (maternal age, gestational age, gravidity, parity); symptoms (headache, visual disturbances, chest pain/ dyspnoea, right upper quadrant pain or epigastric pain, nausea, vomiting, and vaginal bleeding with abdominal pain); signs (blood pressure, and dipstick proteinuria); and laboratory variables (haematocrit, platelet count, white blood cell count, serum creatinine, bilirubin, aspartate aminotransferase (AST), alanine aminotransferase (ALT), lactate dehydrogenase (LDH), international normalized ratio (INR), fibrinogen, albumin, urinary protein: creatinine ratio and

24 hour urinary protein concentration). Only demographics symptoms and signs were considered during miniPIERS model building, as it was designed for use in low-resourced settings were laboratory tests are rarely available. The fullPIERS model considered demographics, symptoms, signs and laboratory variables as candidate predictors.

Data were abstracted by research staff from the medical record. Occurrence of any symptoms and the 'worst' (either highest or lowest depending on the variable) signs were recorded daily while all laboratory values for each day, even when multiple tests were performed on a given day of admission were recorded. For this study, the data on the worst predictor variables collected within 24 hours of hospital admission was considered for model development and validation. For any instance of missing laboratory data within the first 24 hours of admission, the worst value recorded within two weeks prior to admission was used. In the fullPIERS dataset, missing values for SpO_2 were imputed with a value of 97%, the median of the available measures. This was done after interim analysis demonstrated that SpO_2 was highly significant on univariate logistic regression analysis and because missingness could be considered missing at random due to a slow uptake of nursing staff recording the SpO_2 result in the medical record during the first six months of study [119].

6.2.3 Main outcome measures

The primary outcome for both the fullPIERS and miniPIERS studies was a composite maternal outcome including maternal death or major morbidity, as previously described in Chapter 3. Women were considered to have met the requirements for an outcome if one or more of the

components of the composite occurred within 48 hours of admission to hospital. A full list of composite outcome components is provided, with definitions, in Appendix B.

6.2.4 Statistical analysis

Initial Predictor Selection

Prior to model development, predictors were first examined for collinearity. This step was excluded from subsequent automated model building steps because it was decided a priori that in the event of collinearity, variable selection would be based on substantive knowledge rather than statistical criteria. Correlation between variables was determined in a pairwise manner and only the more clinically relevant variable of a pair of highly correlated variables was retained. Clinical relevance was determined through discussion with the study investigators, P von Dadelszen and L Magee. When a high degree of correlation existed between two symptoms ($r>0.5$), they were re-coded as a combined indicator variable. This initial parameter selection step was completed for each dataset for this study to ensure consistent methodology was used rather than relying on the published models candidate predictor list which, for the published fullPIERS model[119], was generated using a cluster analysis technique. The resultant candidate predictor list was then subjected to the remaining model building process in a fully automated fashion.

Model Building

Prediction model development was completed as follows:

Variable coding: Variables with a skewed distribution were assessed using a log-transformation (natural log). Continuous variables were assessed for non-linearity, and were considered as restricted cubic splines (with either 3 or 4 knots) or squared terms when appropriate [107]. Choice of final variable to be included in subsequent model development steps was based on the Akaike information criterion (AIC) and was automated during the model development process. In all cases, the linear term was also considered.

(1) Variable selection: Step-wise backward elimination using a stopping rule of p>0.2 for association of individual variables within the multivariate model was used to select predictors for inclusion in the final model.
(2) Parameter estimation: Maximum likelihood estimation logistic regression was used to estimate the direction and magnitude of the association between each predictor and the primary outcome.

An initial model was developed on the original dataset following the three steps outlined above. Bootstrapping was then performed in three cycles using 200 bootstrap sampling iterations for each cycle. In the first cycle, bootstrap models were developed including all three steps described below. In the second cycle, only the variable selection and parameter estimation steps were used to generate bootstrap models. Finally, in the third cycle, parameter

estimation alone was included in the model development process. In all cycles, model building steps were automated to ensure each step could be easily replicated.

Estimating Model Performance and Optimism

Model optimism was assessed using Efron's enhanced bootstrap procedure applied three ways to each dataset as described above [114]. The bootstrapping procedure involves (1) sampling with replacement from the original cohort to generate a bootstrap dataset of the same size; (2) redevelopment of the model including all model development steps; (3) estimation of the AUC ROC for the model in the bootstrap sample; (4) application of this new model to the original dataset and estimation of AUC ROC. Model optimism is then calculated as the average difference between model performance in the bootstrap sample and the original dataset after 200 iterations of this procedure. Optimism was estimated for: (1) all model development steps (1-3 under the model building subsection above); (2) variable selection and parameter estimation only; and (3) parameter estimation only.

For each model developed, variability in transformation selection was also tabulated.

All statistical analyses were performed using STATA v11·0 (StataCorp, College Station, TX, USA).

6.3 Results

miniPIERS

During the initial parameter selection step, prior to automated model development, strong correlation (r>0·5) was demonstrated between the symptoms of chest pain and dyspnoea, and

headache and visual disturbances. Therefore, these symptoms were re-coded as combined indicator variables and entered accordingly into the multivariate model. As expected, systolic and diastolic blood pressure were highly correlated. Systolic blood pressure was selected for final model development because it is easier for minimally-trained health care providers to measure by radial artery palpation than detection of Korotokoff sounds and has been shown to be reflective of stroke risk in women with pre-eclampsia [76].

Log transformations were tested for both systolic blood pressure measurements and gestational age at admission due to the highly skewed distribution of both variables (Figure 3-1). Non-linear transformations were also tested for systolic blood pressure and gestational age during final model development after plotting as restricted cubic splines revealed non-linear relationships with the outcome in univariate analysis (Figure 3-2).

Table 6-1 presents the proportion of bootstrap models selecting each potential variable transformation for inclusion in the variable selection process. In all cases the transformations selected for any given variable were highly inconsistent between bootstrap samples. For example, prior to backward selection, 36.0% of models considered the linear systolic blood pressure term whereas 31.0%, 16.0% and 15.0% considered the log transformed, 3 knot and 4 knot spline transformations, respectively. The distribution of transformations considered in models was similar for gestational age at admission with 32.5%, 11.0%, 10.0%, 41.0% and 5.5% of models considering the linear, log transformed, 3 knot spline, 4 knot spline and squared terms, respectively. No transformation was clearly favored throughout the bootstrap samples.

As expected, there was also significant variability in parameters chosen for each bootstrap model based on the backward elimination process. The most often selected parameters were the symptom complexes of visual disturbances/ headache and abdominal pain with vaginal bleeding, which were both included in 100% of bootstrap models and the least often selected parameter was nausea, which was selected in 26% of models. Table 6-2 presents the breakdown of variable selection for all bootstrap models. The most common bootstrap model included all parameters except nausea and maternal age and was generated in 56.0% of the bootstrap samples.

Some variability in parameter estimates was observed in the final bootstrap analysis including only the estimation step of model development. The median and interquartile range of the point estimate for all coefficients of variables tested are presented in Figure 6-1 as box plots with whiskers presenting the absolute range. An inconsistent direction and magnitude of effect was found for the predictor variables gestational age at admission, parity, right upper quadrant pain, and dipstick proteinuria. The variable with the greatest variance (0.35) was gestational age at onset (median -1.14 [IQR -1.57, -0.76]).

The AUC ROC for the original model (apparent AUC ROC) was 0.75 (95% CI 0.72 - 0.79). The average optimism found after applying the bootstrap process: (a) including all model development steps of variable coding, variable selection and parameter estimation was 0.034 [95% CI 0.026 - 0.042]; (b) including variable selection and parameter estimation steps only was 0.014 [95% CI 0.011 – 0.015]; and (c) when including parameter estimation only was 0.012 [95%

CI 0.009, 0.014]. The range in AUC ROCs for each bootstrap model in the bootstrap sample and when applied to the original dataset is presented in Figure 6-2a-c.

fullPIERS

Initial variable inspection using the fullPIERS dataset revealed no correlation between candidate predictor symptoms, although unlike miniPIERS, the symptoms of chest pain and dyspnoea were collected as a composite during this study so were already grouped prior to analysis. Laboratory and clinical sign variables that showed a high degree of correlation were: AST, ALT and LDH; systolic and diastolic blood pressure; uric acid and serum creatinine; and protein:creatinine ratio and dipstick proteinuria. After review of the literature and discussion with study investigators (P von Dadelszen and L Magee), the following of the correlated variables were chosen as the most clinically relevant for model development: AST, systolic blood pressure, uric acid and dipstick proteinuria. Other variables to test in the model development process based on our previous literature review [6] were gestational age on admission, parity, all symptoms, albumin, bilirubin and SpO_2. Gestational age, systolic blood pressure and AST were found to be highly skewed and were investigated for model improvement using a log transformation, while platelets, bilirubin, albumin, systolic blood pressure and gestational showed evidence of potential non-linearity in the predictor outcome relationship and were therefore assessed in the modeling process as restricted cubic splines.

Table 6-1 presents the proportion of bootstrap models selecting each potential variable transformation for inclusion in the variable selection process. As with miniPIERS, the

transformations selected for any given variable were highly inconsistent between bootstrap samples. For example, in the case of systolic blood pressure the most often chosen form of the variable was the linear term, considered in 39.5% of bootstrap models. The remaining 32.0%, 20.0% and 8.5% of bootstrap models considered the log transformed, 3 knot or 4 knot spline transformations of systolic blood pressure, respectively. There was no clearly favoured transformation for any variable tested in the bootstrap process. The only exception to this was platelet count, as 100% of models included a 4-knot spline transformation of this variable.

As with the miniPIERS model, some variability was seen in parameter estimates when this step alone was included in the bootstrap process. Figure 6-3 presents box-plots based on the median and interquartile range of point estimates found for all beta coefficients of model parameters. The estimates of coefficients for chest pain/ dyspnea, right upper quadrant pain, albumin, AST, and dipstick all ranged between a positive and negative direction of effect. The greatest variance (0.23) was found in the estimates for chest pain/ dyspnea (median 1.07 [IQR 0.75, 1.34]).

The AUCROC for the original model was 0.85 (95% CI 0.80 - 0.90). The average optimism found after applying the bootstrap process including (a) all model development steps (optimism=0.055 [95% CI 0.052 - 0.057]); (b) variable selection and parameter estimation only (optimism=0.034 [95% CI 0.033 - 0.036]); and (c) parameter estimation only (optimism=0.022 [95% CI 0.021 - 0.023]). The range in AUC ROC for each bootstrap model in the bootstrap sample and when applied to the original dataset is presented (Figure 6-4a-c).

6.4 Discussion

6.4.1 Main findings

In this study, we demonstrated that overfitting is introduced at all stages of risk prediction model development: (1) variable coding and transformation selection, (2) variable selection and (3) parameter estimation. A significant increase in the estimated optimism of both models was demonstrated when the variable coding step was included in the bootstrap model development procedure compared with inclusion of variable selection and parameter estimation or parameter estimation alone.

The selection of candidate predictor variable transformations was highly unstable across bootstrap samples. The number of options for transformations tested ranged from 2-4, in addition to the linear term. These options included a log term, if the variable had demonstrated a skewed distribution, and a square term or 3 or 4 knot spline if the variable appeared to be non-linear in its effect on the outcome in the development dataset. In most cases the transformation selected in each bootstrap sample was split with a minority (approximately 40-50%) of samples at most including any particular transformation. The only variable where a clear non-linear pattern of effect of the parameter on outcome was shown is platelet count in the fullPIERS dataset, as 100% of models included a 4-knot spline transformation of this variable.

Backward selection also introduced significant variability to the final models. In the miniPIERS dataset, the most common model to be generated occurred in 56.0% of bootstrap samples

whereas in the fullPIERS dataset the most common model occurred only 26.0% of the time. This likely reflects the larger number of parameters tested in the fullPIERS model development process, which leaves greater room for uncertainty in the parameter selection step.

6.4.2 Strengths and limitations

The ability to replicate findings using two distinct datasets including women with an HDP is a significant strength of this study. By assessing the effect of all modeling steps in two datasets of similar yet distinct populations we can begin to build hypotheses on the generalizability of these findings to model building overall. The datasets used were both rigorously collected using consistent procedures and definitions and include a large number of cases and outcomes. This allows for comparison between the results and supports the development of the study hypothesis.

Despite the large size of the datasets used and significant occurrence of the adverse maternal outcome, in the fullPIERS dataset in particular, it is likely some overfitting can be attributed to sample size limitations and the number of degrees of freedom tested through the automated variable and transformation selection processes. For the fullPIERS model we tested 15 candidate predictor variables, five of which were assessed using a 4-knot spline and one of which was a categorical variable with four indicator categories resulting in a total of 29 degrees of freedom or effective variables tested. For a model of this size, the dataset should ideally have a minimum of 290 outcomes. The fullPIERS dataset was far short of that with only 106 adverse maternal outcomes occurring. Although this is likely to have introduced additional bias

in our parameter estimates, including all candidate predictors as done in this study is a pragmatic approach as these are the variables that would be tested should we have wanted to develop a model on this dataset for external use. In addition, the bootstrap process was chosen as a way of assessing overfitting in this practical example due to the fact that it best accounts for the bias introduced when using a small sample size [112].

A further limitation of this study is that we were unable to confirm the applicability of these findings to the external validity of the developed models. Assessment of overfitting and internal validity of the risk model is done to understand the likely decrease in model performance when it is applied to a new population. Strengthening our ability to estimate this decrease in performance, or model optimism, by including all model development steps in the bootstrapping process is only relevant if the true performance in an external dataset was more accurately reflected by the higher optimism estimated. Confirmation of our hypothesis that analysts should include all model development steps in the bootstrapping process still requires assessment of the effect of these models in external datasets.

Finally, in this study we chose to test the impact of the various model development steps on model optimism estimated using bootstrapping alone. There are several other techniques commonly used to estimate optimism that we have not assessed, such as leave-1-out-cross-validation, cross-validation with replication and leave-pair-out cross validation. Bootstrapping was chosen for this study as it has been demonstrated to result in the least biased estimates of optimism in simulated datasets [111-113,116,126,159], although one study using both clinical and

simulated datasets did show that cross-validation with replication or leave-pair-out cross validation may give equivalently unbiased results when the events per variable is below the recommended 10 and the data set is small [160]. Cross-validation techniques continue to be used for assessment of internal validity of prediction models and estimation of model optimism [127]. In addition, prediction models are routinely developed using small datasets with less than 10 events per effective variable, so it would be useful to assess the generalizability of this study's findings when using these other cross-validation techniques to estimate model optimism.

6.4.3 Interpretation

Despite the demonstrated accuracy of the bootstrap method to estimate optimism in model performance measures that has been shown in simulation studies [107,111-113], and more recently in a study using clinical datasets[160], many prediction models still perform more poorly than expected on external validation. Effort has been made to understand the reasons for this, which may include differences in case-mix in the external population compared with the model development population, errors in estimation of model parameters during model development or underestimation of model optimism due to use of inaccurate methods of assessing overfitting such as a split sample or cross-validation without replication [108,118,160]. Several studies have addressed the instability of models developed using variable selection methods such as backward elimination and the significant contribution this automated selection process has on model optimism [115,134,161]. We were able to further support these studies conclusions by demonstrating that backward selection did result in unstable parameter selection in the

bootstrap samples and contributed to model overfitting. Through this study we have tested an additional hypothesis that lack of inclusion of all model development steps, including variable coding, in the bootstrap procedure may also contribute to underestimates of model optimism and hence a greater decline in model performance in external datasets than had been expected.

In Steyerberg's book *Clinical Prediction Models: A Practical approach to development, validation and updating,* he outlines sources of potential bias in model performance to include "specification of the structure of our model, such as which characteristics are included as predictors, on information of the data set under study". He then goes on to include selection of variable transformations based on study data as one source of this model uncertainty and bias [111]. Harrell, in his description of the model development and validation process also speculates that there is a potential for overfitting due to both variable selection and assessments of linearity of variables based on the data under study [107]. Despite this speculation of effect, descriptions of the bootstrap method by both Steyerberg and Harrell have suggested inclusion of these early model development steps in the bootstrapping process is unnecessary due to the perceived difficulty in automating these steps [107,111]. This has led to a lack of attention to the effect of this model development step as a source of overfitting and as a result, no study has previously characterized the effect of the variable coding and transformation selection process on estimates of model optimism.

Results from this study suggest that including all model development steps in the bootstrap analysis may produce more reliable estimates of model optimism and allow researchers to have a more reasonable expectation of model performance in new populations. We have demonstrated that there is a significant difference in estimates of model optimism when the variable coding process is included in the bootstrap analysis, as compared to including variable selection and parameter estimation only. This conclusion is further supported by our finding that selection of transformations was highly variable and unstable across the 200 iterations of bootstrap model development.

6.5 Conclusion

Bootstrapping procedures to estimate model performance and optimism of clinical prediction models should include variable coding as well as variable selection and parameter estimation to capture all aspects of overfitting. Further assessment of these findings is required to confirm generalizability of results and relationship to model external validation.

Table 6-1: Distribution of variable transformations selected for inclusion in backward selection process for bootstrap models generated using both the miniPIERS and fullPIERS datasets.

Transformations chosen	miniPIERS (% models)	fullPIERS (% models)
Gestational Age on Admission		
Linear	32.5%	32.5%
Log	11.0%	40.0%
3 knot spline	10.0%	12.5%
4 knot spline	41.0%	14.0%
Square	5.5%	1.5%
Systolic Blood Pressure		
Linear	36.0%	39.5%
Log	31.0%	32.0%
3 knot spline	16.0%	20.0%
4 knot spline	15.0%	8.5%
AST		
Linear	n/a	26.5%

Transformations chosen	miniPIERS (% models)	fullPIERS (% models)
Log	n/a	73.5%
Albumin		
Linear	n/a	22.0%
3 knot spline	n/a	45.0%
4 knot spline	n/a	33.0%
Platelets		
Linear	n/a	0%
3 knot spline	n/a	0%
4 knot spline	n/a	100%
Bilirubin		
Linear	n/a	41.5%
3 knot spline	n/a	52.0%
4 knot spline	n/a	6.5%

Table 6-2: Proportion of bootstrap models selecting each candidate predictor variable

Variable	miniPIERS (% models selected)	fullPIERS (% models selected)
Gestational age on admission	80.0%	32.0%
Parity	78.0%	28.0%
Maternal age	28.0%	n/a
Visual disturbances*/ headache	100.0%	28.0%
Headache*	n/a	32.0%
Chest pain/ dyspnoea	98.0%	90.0%
Abdominal pain with vaginal bleeding	100.0%	100.0%
Right upper quadrant pain	78.0%	64.0%
Nausea	26.0%	22.0%
Systolic blood pressure	94.0%	36.0%
Dipstick proteinuria	86.0%	80.0%
Aspartate transaminase	n/a	58.0%
Albumin	n/a	60.0%
Platelets	n/a	100.0%
Bilirubin	n/a	52.0%
Uric acid	n/a	26.0%
SpO_2	n/a	100.0%

* in the miniPIERS cohort visual disturbances and headache were combined into one indicator variable whereas in the fullPIERS cohort they were assessed independently.

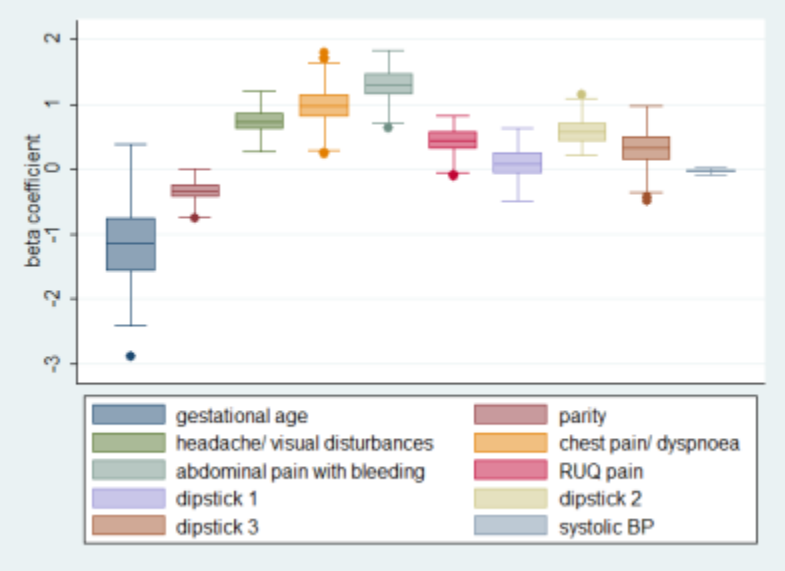


Figure 6-1 Distribution of parameters estimates for variables included in the apparent miniPIERS model

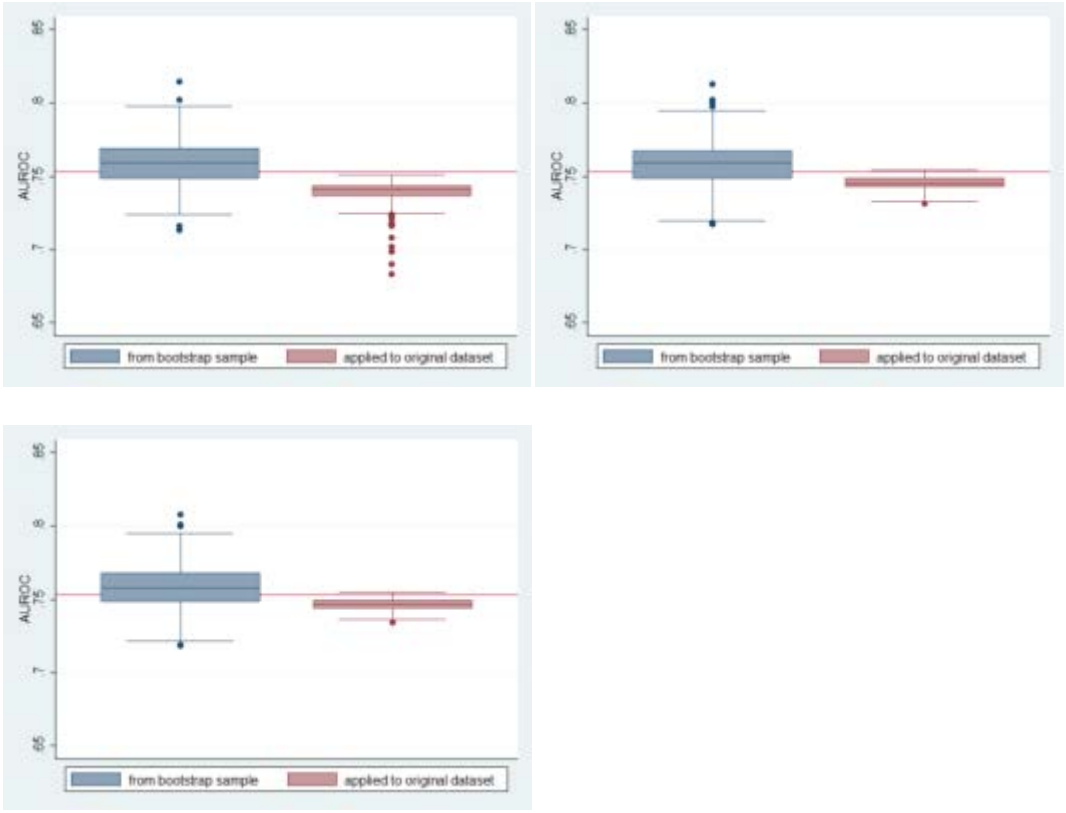


Figure 6-2: Distribution of Area under the receiver operating characteristic curves estimated for models built using the miniPIERS bootstrap samples (blue) and applied to the original miniPIERS data (red). The red horizontal line represents the AUC ROC of the original model. Box plots present median and interquartile range over 200 iterations of bootstrap analysis including (a) all model development steps (optimism=0.034 [95% CI 0.026, 0.042]); (b) variable selection and parameter estimation only (optimism=0.014 [95% CI 0.011, 0.015]); and (c) parameter estimation only (optimism=0.012 [95% CI 0.009, 0.014]).

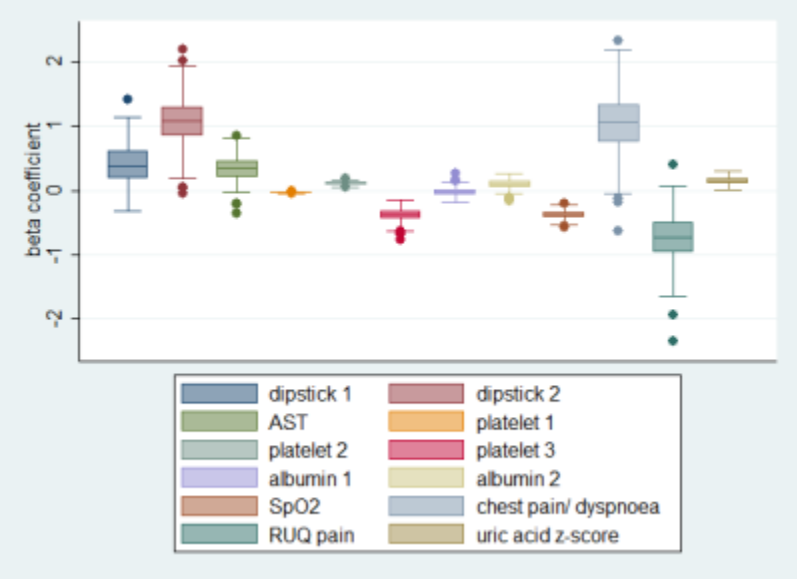


Figure 6-3 Distribution of parameter estimates for variables included in the apparent fullPIERS model

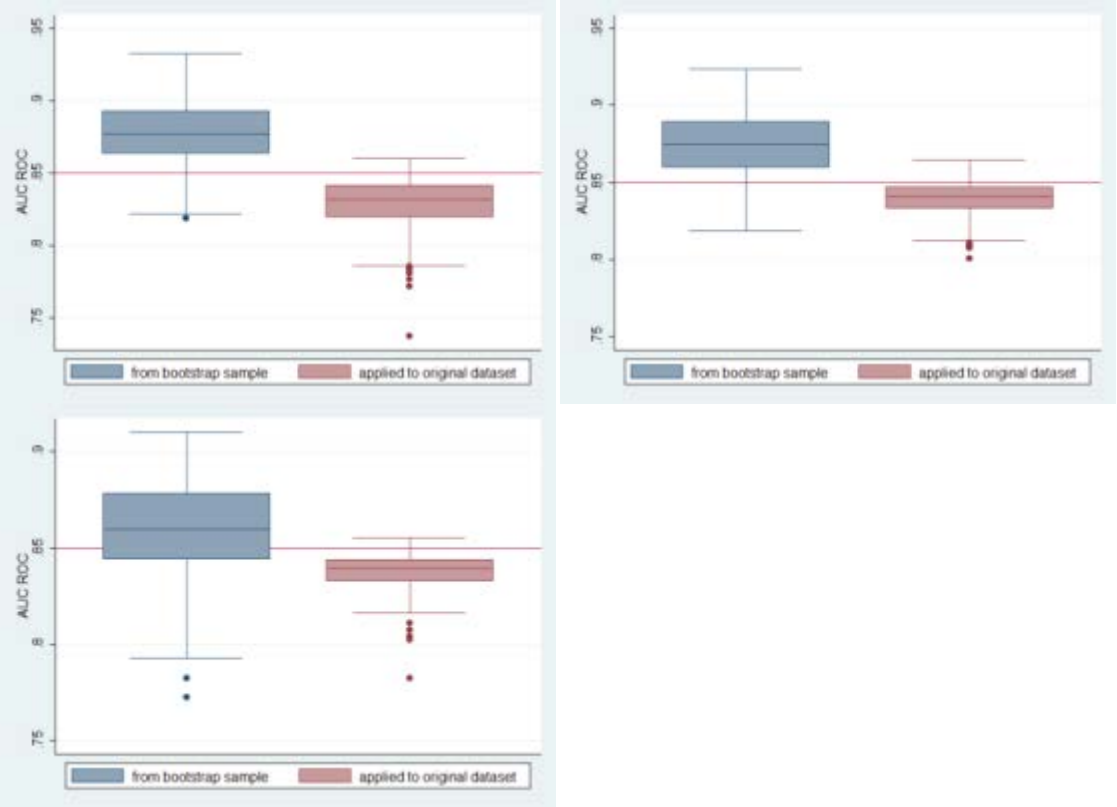


Figure 6-4: Distribution of Area under the receiver operating characteristic curves estimated for models built using the fullPIERS bootstrap samples (blue) and applied to the original fullPIERS data (red). The red horizontal line represents the AUC ROC of the original model. Box plots present median and interquartile range over 200 iterations of bootstrap analysis including (a) all model development steps (optimism=0.055 [95% CI 0.052, 0.057]); (b) variable selection and parameter estimation only (optimism=0.034 [95% CI 0.033, 0.036]); and (c) parameter estimation only (optimism=0.022 [95% CI 0.021, 0.023]).

Chapter 7: Conclusion

The objective of my research was to develop an evidence-based clinical decision support tool the can be used in low-resourced settings by minimally trained health workers to aid in the management of women with a hypertensive disorder of pregnancy (HDP). This objective was based on the hypothesis that simple demographics, symptoms and clinical signs can be used to stratify a population of pregnant women with an HDP into meaningful risk groups as part of a multivariate clinical risk prediction model. In this case, 'meaningful', referred to groups with differential prognosis that can be managed uniquely so that the delays in triage, treatment and transport that result in poor outcomes from the HDP are overcome. In the subsequent chapters, successful development and validation of this tool were described along with an assessment of current methodology for assessing internal validity of clinical risk prediction models. In this chapter the broader impact of these findings on our current understanding of pre-eclampsia management and clinical prediction model methodology is summarized and next steps required to maximize the impact of this research are presented.

7.1 Implications of research findings on management of women with HDP

The process of development and validation of the POM application was multi-staged. The first stage required development of a clinical risk prediction model that allows accurate discrimination and risk stratification of individual women with an HDP identified in a low-resourced setting.

Through the miniPIERS study, we demonstrated that using only simple demographics (gestational age and parity), symptoms (chest pain/ dyspnoea, headache/ visual disturbances, vaginal bleeding with abdominal pain) and signs (blood pressure and dipstick proteinuria), we can accurately and effectively discriminate between women who are and are not at risk of significant complications within 48 hours. Predicting risk within this 48 hour time frame would allow health workers to target high-risk women with effective interventions. These interventions would include referral to a more appropriate facility at which to receive advanced care or the provision of treatments, such as $MgSO_4$, antihypertensives or delivery if the woman is already at an appropriate facility.

Beyond the potential life-saving impact risk stratification in clinical care presents, miniPIERS also contributes significantly to our understanding of disease severity in this population. The results of the miniPIERS study highlight the fact that high blood pressure and proteinuria alone are not the only factors clinicians should be concerned with monitoring. The symptoms and demographic characteristics, such as gestational age, that are included in the miniPIERS model contributed significantly to the model's ability to stratify the population into meaningful risk groups. Addition of SpO_2 into routine antenatal assessments is also warranted, where resources allow, based on the demonstrated improvement to sensitivity of the miniPIERS model resulting from its inclusion. The improved knowledge of disease severity resulting from the miniPIERS study should be used to help guide revision of clinical practice guidelines for management of women with the HDP in resource-constrained settings.

To make this model applicable for use in a community setting and overcome the barrier of lack of access to care that many women in low-resourced settings face due to lack of knowledge, resources or autonomy in decision making [16,17], the miniPIERS model was converted into a decision rule and integrated into a broader decision algorithm with the WHO recommendations for treatment of women with HDP [7]. This presents a significant improvement over available tools to guide management of women with the HDP in these settings. It is the first such evidence-based tool that targets improving care through utilization of existing community health workers. Using a process of iterative review and feedback from relevant stakeholders, we have designed the PIERS on the Move decision algorithm for incorporation into a novel, evidence-based mobile phone application [142]. This application will allow minimally trained community-based health workers to assess women for risk of adverse maternal outcomes associated with the hypertensive disorders of pregnancy and respond to this risk immediately. Use of this tool should improve the health workers' ability to refer women who are truly at risk to higher-level facilities in a timely manner so that these women can receive life-saving care.

Current management of women with HDP in low- and middle- income countries (LMICs) is limited to facility-based care due to a lack of community-based health workers trained to identify and manage this disorder. This presents several obstacles to the effective reduction of outcomes that make the HDP of great concern. First, the majority of women in these settings do not routinely access formal health care but, instead, rely on systems of traditional or lay-health workers to receive antenatal care [16,17]. Secondly, due to the delays in identification and management of the HDP by the existing traditional and lay health worker networks, women

often do not present to care at formal health facilities until it is too late to avoid life-altering morbidity or mortality [51]. To improve outcomes of women whose pregnancies are complicated by the HDP in LMICs it is necessary to improve the ability of community-based health workers to identify women at risk of serious complications of the HDP and get them to life-saving care before it is too late; this is what we have designed the miniPIERS model and PIERS on the Move (POM) application to do. The miniPIERS model and PIERS on the Move tool present an opportunity to overcome the barriers to access of care by providing the community-based health workers an evidence-based tool that allows them to identify and triage high-risk women in a timely manner.

7.2 Strengths and limitations of thesis research

Strength and weaknesses of each individual stage of research presented in this dissertation are described in the individual chapters but there are overarching strengths and weaknesses to this body of work that warrant further discussion here. As outlined in the introduction to this dissertation, task-shifting aspects of pregnancy care to the existing health workforce (in this case community-based health workers) has shown potential to result in improved perinatal survival and was associated with increased health service utilization overall [19,60]. Task-shifting has several recognized risks, for example, potential overburdening of existing health workers and a lack of higher-level health resource infrastructure to support and supervise the expansion of responsibility in lower level cadres. An additional criticism is that many of the efforts to utilize community-based health workers have been research- or subject-led, rather than needs-

led which has resulted in limited potential for sustainability of programmes once non-governmental organization of charitable organization funding runs out [14,19,53,54,64].

One of the strengths of the research presented in this dissertation is that we have designed the miniPIERS model and PIERS on the Move application in close collaboration with a large network of stakeholders and health workers from the LMIC settings in which we hope to implement the tool. Wherever possible we have utilized a needs-led and user-led design process for tool development so that when miniPIERS, through the PIERS on the Move application, is implemented into health systems in LMICs it will be feasible and acceptable for use. More importantly, inclusion of relevant stakeholders such as members of the Ministry of Health and senior clinicians from our target countries in the PIERS on the Move tool design process improves the likelihood of sustainability of an antenatal care programme that includes this tool after research into the impact of implementation of this programme is complete, should results be positive.

Use of a user-centred design process does not address concerns with lack of infrastructure at the higher-level health facility. We have also not been able to address the concern with lack of highly skilled staff to supervise and support task-shifting but by designing the PIERS on the Move tool with target users, we are ensuring the tool is as simple and easy to implement as possible. This will work towards addressing the concern of overburdening the community-based health workers with new complex tasks when introducing a tool such as PIERS on the Move. It is important to note that we have achieved this simplicity while still ensuring high

performance levels are reached for the miniPIERS model. Use of mHealth technology for application of this model will further improve acceptability of the tool as it utilizes technology that is familiar and reliable in the LMIC setting. Further to this, creating a mHealth application further simplifies application of the model, by eliminating the need for the health worker to complete any calculations, and allows it to be available during home visits, overcoming the barrier many women face of access to care at formal health facilities.

The greatest limitation of the research presented here is that, although we demonstrated significant potential to improve management of women with HDP through use of the miniPIERS model and PIERS on the Move tool, we did not demonstrate the magnitude of this impact on clinical outcomes. The study completed in South Africa, as described in Chapter 5, suffered from lower recruitment rates than expected and only utilized the PIERS on the Move application for data collection. Using data from this study we can only estimate expected referral rates and health system impacts in terms of time taken to complete an antenatal assessment. This work provided valuable preliminary information for the design of an implementation study (the CLIP trial) but it is not until this implementation study is complete that the miniPIERS model can be justifiably scaled-up and implemented into routine care.

7.3 Future research directions

The research presented in this dissertation has impact on both clinical care of women with HDP but also increases our knowledge of the impact of methodology used to develop and validate prediction models. An important contribution of this research was the observation that

selection of variable transformations and categorizations based on the data on hand contributes significantly to prediction model overfitting and optimism. This had been suggested in the literature previously [107,162,163] but had not been demonstrated in a practical way. Further work is required to confirm the generalizability of these findings to development of other prediction models and to assess the impact of these findings on external validity of the PIERS models themselves. Generalizability of the results should be tested using an additional dataset in which the full process of model development, internal and external validation can be completed. Further confirmation of findings and conclusions drawn in relation to prediction model methodology overall could also be achieved through a simulation study. Simulating data with known parameters and assumptions would allow us to better characterize the contribution of the various model development steps to model optimism, in a controlled manner. Assessment of external validity of the miniPIERS models is also planned. Once this work is complete we will be able to draw conclusions on the effect of variable coding steps such as transformation selection on overall validity of clinical prediction models. This information would improve other researchers' chance of developing models that are valid and effective and, therefore, can improve clinical care through more accurate prognosis and diagnosis.

The greatest impact of this research will be achieved once the PIERS on the Move tool is applied in clinical practice and tested against impact on both maternal and perinatal outcomes. We are currently testing the impact of the PIERS on the Move tool through the Community Level Interventions for Pre-eclamspia (CLIP) trial [144]. The results of this trial will help guide subsequent policy advocacy work required to scale-up the PIERS on the Move application.

We believe that this clinical decision support tool, PIERS on the Move, is an important contribution as it offers the potential to improve health outcomes of women for a condition that is at the root of a large amount of morbidity and mortality in the developing world. Pre-eclampsia alone accounts for an estimated 18.5% of all maternal death, globally. Looking beyond the impact on maternal health outcomes, there are also societal impacts possible through improved health service delivery with the use of an effective decision support tool such as PIERS on the Move. The potential societal implications of introduction of this tool into routine antenatal care for LMICs are twofold: first, at the individual level women would not suffer the cost and time away from their families for unnecessary referrals when safe, increased community surveillance would be appropriate. Secondly, at the health system level, evidence-based monitoring and primary triage for women with the HDPs (especially pre-eclampsia) is moved from the tertiary facilities alone into lower level or primary health clinics, thereby increasing the potential for broad population-based screening, as well as making more efficient use of already burdened acute care facilities.

Even if the CLIP trial is successful and shows significant positive results, this may not be sufficient to warrant scale-up of the PIERS on the Move tool alone due to its singular focus on the HDP. To improve potential impact of the PIERS on the Move tool, expanding the scope of the tool to introduce risk prediction models addressing the other major causes of maternal mortality would be highly beneficial. Repeating the PIERS process to develop and validate a risk prediction model to better understand prognosis of both postpartum hemorrhage and obstetric sepsis and the impact of various available interventions that can be applied immediately by

community-based health workers in a community setting is a logical next step to the research presented in this dissertation. This expansion of scope would further justify scale-up of the PIERS on the Move tool in a resource constrained setting as it would make better use of the available health resources.

Offering a single integrated mHealth solution, rather than several applications focused on various health conditions is necessary to ensure that these innovations are actually beneficial to the community-based health workers. In doing this research, the biggest risk to sustainability should the CLIP trial have positive results will be a lack of ability to integrate the PIERS on the Move tool into routine practice due to its highly focused content.

Our goal is to support task-shifting aspects of clinical care to available health resources. This can only be achieved if the tools we are creating are simple and enhance the health workers ability to complete tasks assigned. It is important that these tools work together with other applications used by the community-based health workers and the broader health system. One way to ensure successful integration will be to expand the scope of the PIERS on the Move decision support platform to other causes of maternal mortality and convert the tool into a complete platform for guiding antenatal care assessments at the community level. Finally, as we consider how to move forward to have the greatest impact, a continued focus on the opinions and needs of the community-based health workers themselves and the population they serve is required.

References

1. Magee LA, Pels A, Helewa ME, Rey E, von Dadelszen P. Diagnosis, evaluation and management of the hypertensive disorders of pregnancy. *Pregnancy Hypertension*. 2014;4:105-106-145.

2. National Institute for Health and Clinical Excellence. Hypertension in pregnancy: The management of hypertensive disorders during pregnancy (clinical guideline 107). . 2010.

3. Lindheimer MD, Taler SJ, Cunningham FG, American Society of Hypertension. ASH position paper: Hypertension in pregnancy. *J Clin Hypertens (Greenwich)*. 2009;11(4):214-225.

4. Lowe SA, Brown MA, Dekker GA, et al. Guidelines for the management of hypertensive disorders of pregnancy 2008. *Aust N Z J Obstet Gynaecol*. 2009;49(3):242-246.

5. Menzies J, Magee LA, Macnab YC, et al. Current CHS and NHBPEP criteria for severe preeclampsia do not uniformly predict adverse maternal or perinatal outcomes. *Hypertens Pregnancy*. 2007;26(4):447-462.

6. Payne B, Magee LA, von Dadelszen P. Assessment, surveillance and prognosis in pre-eclampsia. *Best Practice & Research in Clinical Obstetrics & Gynaecology*. 2011;25(4):449-462.

7. World Health Organization, Dept. of Reproductive Health and Research, Dept. of Maternal, Newborn, Child and Adolescent Health, Dept. of Nutrition for Health and Development. **WHO**

recommendations for prevention and treatment of pre-eclampsia and eclampsia. http://www.who.int/reproductivehealth/publications/maternal_perinatal_health/9789241548335/en/. Updated 2011. Accessed 07/05, 2012.

8. Hutcheon JA, Lisonkova S, Joseph KS. Epidemiology of pre-eclampsia and the other hypertensive disorders of pregnancy. *Best Pract Res Clin Obstet Gynaecol*. 2011;25(4):391-403.

9. Abalos E, Cuesta C, Grosso AL, Chou D, Say L. Global and regional estimates of preeclampsia and eclampsia: A systematic review. *European Journal of Obstetrics, Gynecology, & Reproductive Biology*. 2013;170(1):1-7.

10. Lozano R, Naghavi M, Foreman K, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: A systematic analysis for the global burden of disease study 2010. *Lancet*. 2013;380(9859):2095-2128.

11. Hogan MC, Foreman KJ, Naghavi M, et al. Maternal mortality for 181 countries, 1980-2008: A systematic analysis of progress towards millennium development goal 5. *Lancet*. 2010;375(9726):1609-1623.

12. Moodley J. Maternal deaths due to hypertensive disorders in pregnancy. *Best Practice & Research in Clinical Obstetrics & Gynaecology*. 2008;22(3):559-567.

13. Khan KS, Wojdyla D, Say L, Gulmezoglu AM, Van Look PF. WHO analysis of causes of maternal death: A systematic review. *Lancet*. 2006;367(9516):1066-1074.

14. Joint Learning Initiative. Human resources for health: Overcoming the crisis. http://www.who.int/hrh/documents/JLi_hrh_report.pdf?ua=1. Updated 2004. Accessed 08/03, 2012.

15. Bhutta ZA, Black RE. Global maternal, newborn, and child health--so near and yet so far. *N Engl J Med*. 2013;369(23):2226-2235.

16. Simkhada B, Teijlingen ER, Porter M, Simkhada P. Factors affecting the utilization of antenatal care in developing countries: Systematic review of the literature. *J Adv Nurs*. 2008;61(3):244-260.

17. Gabrysch S, Campbell OM. Still too far to walk: Literature review of the determinants of delivery service use. *BMC Pregnancy & Childbirth*. 2009;9:34.

18. Thaddeus S, Maine D. Too far to walk: Maternal mortality in context. *Soc Sci Med*. 1994;38(8):1091-1110.

19. Fulton BD, Scheffler RM, Sparkes SP, Auh EY, Vujicic M, Soucat A. Health workforce skill mix and task shifting in low income countries: A review of recent evidence. *Human Resources for Health [Electronic Resource]*. 2011;9(1):1.

20. Staff AC, Benton SJ, von Dadelszen P, et al. Redefining preeclampsia using placenta-derived biomarkers. *Hypertension*. 2013;61(5):932-942.

21. Steegers EA, von Dadelszen P, Duvekot JJ, Pijnenborg R. Pre-eclampsia. *Lancet*. 2010;376(9741):631-644.

22. Silasi M, Cohen B, Karumanchi SA, Rana S. Abnormal placentation, angiogenic factors, and the pathogenesis of preeclampsia. *Obstet Gynecol Clin North Am*. 2010;37(2):239-253.

23. Gruslin A, Lemyre B. Pre-eclampsia: Fetal assessment and neonatal outcomes. *Best Pract Res Clin Obstet Gynaecol*. 2011;25(4):491-507.

24. Jim B, Sharma S, Kebede T, Acharya A. Hypertension in pregnancy: A comprehensive update. *Cardiol Rev*. 2010;18(4):178-189.

25. Barton JR, O'brien JM, Bergauer NK, Jacques DL, Sibai BM. Mild gestational hypertension remote from term: Progression and outcome. *Am J Obstet Gynecol*. 2001;184(5):979-983.

26. Homer CS, Brown MA, Mangos G, Davis GK. Non-proteinuric pre-eclampsia: A novel risk indicator in women with gestational hypertension. *J Hypertens*. 2008;26(2):295-302.

27. Saudan P, Brown MA, Buddle ML, Jones M. Does gestational hypertension become pre-eclampsia? *Br J Obstet Gynaecol*. 1998;105(11):1177-1184.

28. Sibai BM, Stella CL. Diagnosis and management of atypical preeclampsia-eclampsia. *Am J Obstet Gynecol*. 2009;200(5):481.e1-481.e7.

29. Haddad B, Barton JR, Livingston JC, Chahine R, Sibai BM. Risk factors for adverse maternal outcomes among women with HELLP (hemolysis, elevated liver enzymes, and low platelet count) syndrome. *Am J Obstet Gynecol*. 2000;183(2):444-448.

30. Magee LA, von Dadelszen P, Chan S, et al. The control of hypertension in pregnancy study pilot trial. *BJOG*. 2007;114(6):770, e13-20.

31. Tuffnell DJ, Jankowicz D, Lindow SW, et al. Outcomes of severe pre-eclampsia/eclampsia in yorkshire 1999/2003. *BJOG*. 2005;112(7):875-880.

32. Abalos E, Cuesta C, Carroli G, et al. Pre-eclampsia, eclampsia and adverse maternal and perinatal outcomes: A secondary analysis of the world health organization multicountry survey on maternal and newborn health. *BJOG: An International Journal of Obstetrics & Gynaecology*. 2014;121(Suppl 1):14-24.

33. Magee LA, von Dadelszen P. The management of severe hypertension. *Semin Perinatol*. 2009;33(3):138-142.

34. von Dadelszen P, Magee LA. Fall in mean arterial pressure and fetal growth restriction in pregnancy hypertension: An updated metaregression analysis. *J Obstet Gynaecol Can*. 2002;24(12):941-945.

35. Duley L, Farrell B, Spark P, et al. Do women with pre-eclampsia, and their babies, benefit from magnesium sulphate? the magpie trial: A randomised placebo-controlled trial. *Lancet*. 2002;359(9321):1877-1890.

36. Gordon R, Magee LA, Payne B, et al. Magnesium sulphate for the management of preeclampsia and eclampsia in low and middle income countries: A systematic review of tested dosing regimens. *Journal of Obstetrics & Gynaecology Canada: JOGC*. 2014;36(2):154-163.

37. Khashu M, Narayanan M, Bhargava S, Osiovich H. Perinatal outcomes associated with preterm birth at 33 to 36 weeks' gestation: A population-based cohort study. *Pediatrics*. 2009;123(1):109-113.

38. McLaurin KK, Hall CB, Jackson EA, Owens OV, Mahadevia PJ. Persistence of morbidity and cost differences between late-preterm and term infants during the first year of life. *Pediatrics*. 2009;123(2):653-659.

39. Haddad B, Sibai BM. Expectant management in pregnancies with severe pre-eclampsia. *Semin Perinatol*. 2009;33(3):143-151.

40. Ganzevoort W, Rep A, de Vries JI, Bonsel GJ, Wolf H, PETRA-investigators. Prediction of maternal complications and adverse infant outcome at admission for temporizing management of early-onset severe hypertensive disorders of pregnancy. *Am J Obstet Gynecol*. 2006;195(2):495-503.

41. Ganzevoort W, Sibai BM. Temporising versus interventionist management (preterm and at term). *Best Practice & Research in Clinical Obstetrics & Gynaecology*. 2011;25(4):463-476.

42. Magee LA, Yong PJ, Espinosa V, Cote AM, Chen I, von Dadelszen P. Expectant management of severe preeclampsia remote from term: A structured systematic review. *Hypertens Pregnancy*. 2009;28(3):312-347.

43. Bombrys AE, Barton JR, Habli M, Sibai BM. Expectant management of severe preeclampsia at 27(0/7) to 33(6/7) weeks' gestation: Maternal and perinatal outcomes according to gestational age by weeks at onset of expectant management. *Am J Perinatol*. 2009;26(6):441-446.

44. Bombrys AE, Barton JR, Nowacki EA, et al. Expectant management of severe preeclampsia at less than 27 weeks' gestation: Maternal and perinatal outcomes according to gestational age by weeks at onset of expectant management. *Am J Obstet Gynecol*. 2008;199(3):247.e1-247.e6.

45. Hall DR, Odendaal HJ, Steyn DW. Expectant management of severe pre-eclampsia in the mid-trimester. *European Journal of Obstetrics Gynecology and Reproductive Biology*. 2001;96(2):168-172.

46. Hall DR, Odendaal HJ, Kirsten GF, Smith J, Grove D. Expectant management of early onset, severe pre-eclampsia: Perinatal outcome. *Br J Obstet Gynaecol*. 2000;107(10):1258-1264.

47. Hall DR, Odendaal HJ, Steyn DW, Grove D. Expectant management of early onset, severe pre-eclampsia: Maternal outcome. *Br J Obstet Gynaecol*. 2000;107(10):1252-1257.

48. Koopmans CM, Bijlenga D, Groen H, et al. Induction of labour versus expectant monitoring for gestational hypertension or mild pre-eclampsia after 36 weeks' gestation (HYPITAT): A multicentre, open-label randomised controlled trial. *Lancet*. 2009;374(9694):979-988.

49. Hutcheon JA, Lisonkova S, Magee LA, et al. Optimal timing of delivery in pregnancies with pre-existing hypertension. *BJOG*. 2011;118(1):49-54.

50. Abdel-Hady E, Fawzy M, El-Negeri M, Nezar M, Ragab A, Helal AS. Is expectant management of early-onset severe preeclampsia worthwhile in low-resource settings? *Arch Gynecol Obstet*. 2010;282(1):23-27.

51. Firoz T, Sanghvi H, Merialdi M, von Dadelszen P. Pre-eclampsia in low and middle income countries. *Best Practice & Research in Clinical Obstetrics & Gynaecology*. 2011;25(4):537-548.

52. Pakenham-Walsh N. Improving access to reliable information in developing countries. *BMJ*. 2000;321(7264):831.

53. Pakenham-Walsh N. Health in an unequal world: What about the needs of health workers in developing countries?. *BMJ*. 2007;335(7627):954.

54. Pakenham-Walsh N. Towards a collective understanding of the information needs of health care providers in low-income countries, and how to meet them. *J Health Commun*. 2012;17(Suppl 2):9-17.

55. Pakenham-Walsh NM. Millennium development goals. focus on the needs of the existing health workforce. *BMJ*. 2010;341:5466.

56. World Health Organization. Optimizing health worker roles to improve access to key maternal and newborn health interventions through task shifting. http://apps.who.int/iris/bitstream/10665/77764/1/9789241504843_eng.pdf. Updated 2012. Accessed 03, 2014.

57. Wall SN, Lee AC, Carlo W, et al. Reducing intrapartum-related neonatal deaths in low- and middle-income countries-what works? *Semin Perinatol*. 2010;34(6):395-407.

58. Lassi ZS, Haider BA, Bhutta ZA. Community-based intervention packages for reducing maternal and neonatal morbidity and mortality and improving neonatal outcomes. *Cochrane Database of Systematic Reviews*. 2010(11):007754.

59. Haws RA, Yakoob MY, Soomro T, Menezes EV, Darmstadt GL, Bhutta ZA. Reducing stillbirths: Screening and monitoring during pregnancy and labour. *BMC Pregnancy Childbirth*. 2009;9 Suppl 1:S5.

60. Lewin S, MunabiBabigumira S, Glenton C, et al. Lay health workers in primary and community health care for maternal and child health and the management of infectious diseases. *Cochrane Database of Systematic Reviews*. 2010;3.

61. Glenton C, Colvin CJ, Carlsen B, et al. Barriers and facilitators to the implementation of lay health worker programmes to improve access to maternal and child health: Qualitative evidence synthesis. *Cochrane Database of Systematic Reviews*. 2013;12.

62. Owolabi OO, Glenton C, Lewin S, Pakenham-Walsh N. Stakeholder views on the incorporation of traditional birth attendants into the formal health systems of low-and middle-income countries: A qualitative analysis of the HIFA2015 and CHILD2015 email discussion forums. *BMC Pregnancy & Childbirth*. 2014;14:118.

63. World Health Organization. Global health workforce alliance and WHO: *Global experience of community health workers for delivery of health related millennium development goals: A systematic review, country case studies, and recommendations for integration into national health systems*. . http://www.who.int/workforcealliance/knowledge/publications/alliance/CHWreport_exsummary.pdf. Updated 2010. Accessed 05/16, 2011.

64. Narasimhan V, Brown H, Pablos-Mendez A, et al. Responding to the global human resources crisis. *Lancet*. 2004;363(9419):1469-1472.

65. Pakenham-Walsh N, Bukachi F. Information needs of health care workers in developing countries: A literature review with a focus on africa. *Human Resources for Health [Electronic Resource]*. 2009;7:30.

66. Sullivan TM, Harlan SV, Pakenham-Walsh N, Ouma S. Working together to meet the information needs of health care providers, program managers, and policy makers in low- and middle-income countries. *J Health Commun*. 2012;17(Suppl 2):5-8.

67. Pakenham-Walsh N. Dying for knowledge. *CMAJ Canadian Medical Association Journal*. 2009;181(6-7):399.

68. mHealth Alliance P, W.C. mHealth and MNCH: State of the evidence. http://mhealthalliance.org/images/content/un_007_evidencegapreport_digital_aaa.pdf. Updated 2013. Accessed 02, 2013.

69. Tamrat T, Kachnowski S. Special delivery: An analysis of mHealth in maternal and newborn health programs and their outcomes around the world. *Maternal & Child Health Journal*. 2012;16(5):1092-1101.

70. International Telecommunication Union. ICT factsheet 2013. http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2013-e.pdf. Updated 2013. Accessed 02/15, 2014.

71. Braun R, Catalani C, Wimbush J, Israelski D. Community health workers and mobile technology: A systematic review of the literature. *PLoS ONE [Electronic Resource]*. 2013;8(6):e65772.

72. Noordam AC, Kuepper BM, Stekelenburg J, Milen A. Improvement of maternal health services through the use of mobile phones. *Tropical Medicine & International Health*. 2011;16(5):622-626.

73. Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: Application and impact of prognostic models in clinical practice. *BMJ*. 2009;338:b606.

74. Moons KGM, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: What, why, and how? *Br Med J*. 2009;338:b375.

75. Hansebout RR, Cornacchi SD, Haines T, Goldsmith CH, Evidence-Based Surg Working Grp. How to use an article about prognosis. *Canadian Journal of Surgery*. 2009;52(4):328-336.

76. Martin JN,Jr, Thigpen BD, Moore RC, Rose CH, Cushman J, May W. Stroke and severe preeclampsia and eclampsia: A paradigm shift focusing on systolic blood pressure. *Obstet Gynecol*. 2005;105(2):246-254.

77. Zhang J, Klebanoff MA, Roberts JM. Prediction of adverse outcomes by common definitions of hypertension in pregnancy. *Obstet Gynecol*. 2001;97(2):261-267.

78. Menzies J, von Dadelszen P, PIERS Study Grp. The PIERS (pre-eclampsia integrated estimate of risk) models: Univariable and cluster analyses. *Hypertension in Pregnancy*. 2008;27(4):620-620.

79. Lindheimer MD, Kanter D. Interpreting abnormal proteinuria in pregnancy: The need for a more pathophysiological approach. *Obstet Gynecol*. 2010;115(2 Pt 1):365-375.

80. Cote AM, Firoz T, Mattman A, Lam EM, von Dadelszen P, Magee LA. The 24-hour urine collection: Gold standard or historical practice? *Am J Obstet Gynecol*. 2008;199(6):625.e1-625.e6.

81. Cote AM, Brown MA, Lam E, et al. Diagnostic accuracy of urinary spot protein:Creatinine ratio for proteinuria in hypertensive pregnant women: Systematic review. *BMJ*. 2008;336(7651):1003-1006.

82. Brown MA, Lindheimer MD, de Swiet M, Van Assche A, Moutquin JM. The classification and diagnosis of the hypertensive disorders of pregnancy: Statement from the international society for the study of hypertension in pregnancy (ISSHP). *Hypertens Pregnancy*. 2001;20(1):IX-XIV.

83. Chan P, Brown M, Simpson JM, Davis G. Proteinuria in pre-eclampsia: How much matters? *Bjog-an International Journal of Obstetrics and Gynaecology*. 2005;112(3):280-285.

84. Payne B, Magee LA, Cote AM, et al. PIERS proteinuria: Relationship with adverse maternal and perinatal outcome. *Journal of Obstetrics & Gynaecology Canada: JOGC*. 2011;33(6):588-597.

85. Thangaratinam S, Coomarasamy A, O'Mahony F, et al. Estimation of proteinuria as a predictor of complications of pre-eclampsia: A systematic review. *BMC Med*. 2009;7:10.

86. Ganzevoort W, Rep A, De Vries JI, Bonsel GJ, Wolf H, PETRA-Investigators. Relationship between thrombophilic disorders and type of severe early-onset hypertensive disorder of pregnancy. *Hypertens Pregnancy*. 2007;26(4):433-445.

87. Valera M, Parant O, Vayssiere C, Arnal J, Payrastre B. Physiologic and pathologic changes of platelets in pregnancy. *Platelets*. 2010;21(8):587-595.

88. Romero R, Mazor M, Lockwood CJ, et al. Clinical significance, prevalence, and natural history of thrombocytopenia in pregnancy-induced hypertension. *Am J Perinatol*. 1989;6(1):32-38.

89. Brown MA, Buddle ML. Hypertension in pregnancy: Maternal and fetal outcomes according to laboratory and clinical features. *Med J Aust*. 1996;165(7):360-365.

90. Laskin S, Payne B, Hutcheon J, von Dadelszen P, PIERS Study Grp. The role of platelet counts in the assessment on inpatient women with pre-eclampsia. *Pregnancy Hypertension*. 2010;1(1):S7.

91. Martin JN,Jr, May WL, Magann EF, Terrone DA, Rinehart BK, Blake PG. Early risk assessment of severe preeclampsia: Admission battery of symptoms and laboratory tests to predict likelihood of subsequent significant maternal morbidity. *Am J Obstet Gynecol*. 1999;180(6 Pt 1):1407-1414.

92. Cavkaytar S, Ugurlu EN, Karaer A, Tapisiz OL, Danisman N. Are clinical symptoms more predictive than laboratory parameters for adverse maternal outcome in HELLP syndrome? *Acta Obstet Gynecol Scand*. 2007;86(6):648-651.

93. Benton SJ, von Dadelszen P, PIERS (Pre-eclampsia Integrated Estimate of RiSk) Study Group. Abnormal liver tests as predictors for adverse maternal outcomes in women with pre-eclampsia. *Pregnancy Hypertension*. 2010;1(1):P88.

94. Martin AC, Brown MA. Could uric acid have a pathogenic role in pre-eclampsia? *Nat Rev Nephrol*. 2010.

95. Bainbridge SA, Roberts JM, von Versen-Hoynck F, Koch J, Edmunds L, Hubel CA. Uric acid attenuates trophoblast invasion and integration into endothelial cell monolayers. *American Journal of Physiology - Cell Physiology*. 2009;297(2):440-450.

96. Bainbridge SA, Roberts JM. Uric acid as a pathogenic factor in preeclampsia. *Placenta*. 2008;29(Suppl A):S67-72.

97. Williams KP, Galerneau F. The role of serum uric acid as a prognostic indicator of the severity of maternal and fetal complications in hypertensive pregnancies. *J Obstet Gynaecol Can*. 2002;24(8):628-632.

98. Thangaratinam S, Ismail KMK, Sharp S, Coomarasamy A, Khan KS, TIPPS review grp. Accuracy of serum uric acid in predicting complications of pre-eclampsia: A systematic review. *Bjog-an International Journal of Obstetrics and Gynaecology*. 2006;113(4):369-378.

99. Parrish M, Griffin M, Morris R, Darby M, Owens MY, Martin JN. Hyperuricemia facilitates the prediction of maternal and perinatal adverse outcome in patients with severe/superimposed preeclampsia. *J Matern Fetal Neonatal Med*. 2010.

100. Millman A, Payne B, von Dadelszen P, PIERS (Pre-eclampsia Integrated Estimate of RiSk) Study Group. Oxygen saturation as a predictor of outcomes in women with pre-eclampsia. *Pregnancy Hypertension*. 2010;1(1):S58.

101. Yen TW, Payne B, Qu Z, et al. Using clinical symptoms to predict adverse maternal and perinatal outcomes in women with preeclampsia: Data from the PIERS (pre-eclampsia integrated estimate of RiSk) study. *Journal of Obstetrics & Gynaecology Canada: JOGC*. 2011;33(8):803-809.

102. Kannel WB, Mcgee D, Gordon T. General cardiovascular risk profile - framingham study. *Am J Cardiol*. 1976;38(1):46-51.

103. Nutter AL, Tanawuttiwat T, Silver MA. Evaluation of 6 prognostic models used to calculate mortality rates in elderly heart failure patients with a fatal heart failure admission. *Congestive Heart Failure*. 2010;16(5):196-201.

104. Wong DT, Knaus WA. Predicting outcome in critical care - the current status of the apache prognostic scoring system. *Canadian Journal of Anaesthesia-Journal Canadien D Anesthesie*. 1991;38(3):374-383.

105. Knaus WA, Harrell FE,Jr, Lynn J, et al. The SUPPORT prognostic model. objective estimates of survival for seriously ill hospitalized adults. study to understand prognoses and preferences for outcomes and risks of treatments. *Ann Intern Med*. 1995;122(3):191-203.

106. Richardosn DK, Corcoran JD, Escobar GJ, et al. SNAP-II and SNAPPE-II: Simplified newborn illness severity and mortality risk scores. *J Pediatr*. 2001;138(1):92-100.

107. Harrell FE,Jr, Lee KL, Mark DB. Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996;15(4):361-387.

108. Altman DG, Vergouwe Y, Royston P, Moons KGM. Prognosis and prognostic research: Validating a prognostic model. *Br Med J*. 2009;338:b605.

109. Harrell FE, Lee KL, Pollock BG. Regression-models in clinical-studies - determining relationships between predictors and response. *J Natl Cancer Inst*. 1988;80(15):1198-1202.

110. Picard RR, Berk KN. Data splitting. *The American Statistician*. 1990;44(2):140.

111. Steyerberg EW. *Clinical prediction models: A practical approach to development, validation and updating.* 1st ed. New York: Springer; 2009.

112. Steyerberg EW, Bleeker SE, Moll HA, Grobbee DE, Moons KG. Internal and external validation of predictive models: A simulation study of bias and precision in small samples. *J Clin Epidemiol*. 2003;56(5):441-447.

113. Koller MT, Steyerberg EW. Some methodologic issues in validation of prognostic models. *J Card Fail*. 2006;12(9):759-761.

114. Efron B TR. An Introduction to the Bootstrap. Chapman & Hall. New York. 1993.

115. Steyerberg EW, Eijkemans MJ, Harrell FE,Jr, Habbema JD. Prognostic modelling with logistic regression analysis: A comparison of selection and estimation methods in small data sets. *Stat Med*. 2000;19(8):1059-1079.

116. Vergouwe Y, Steyerberg EW, Eijkemans MJ, Habbema JD. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. *J Clin Epidemiol*. 2005;58(5):475-483.

117. Royston P, Moons KGM, Altman DG, Vergouwe Y. Prognosis and prognostic research: Developing a prognostic model. *Br Med J*. 2009;338:b604.

118. Vergouwe Y, Moons KG, Steyerberg EW. External validity of risk models: Use of benchmark values to disentangle a case-mix effect from incorrect coefficients. *Am J Epidemiol*. 2010;172(8):971-980.

119. von Dadelszen P, Payne B, Li J, et al. Prediction of adverse maternal outcomes in pre-eclampsia: Development and validation of the fullPIERS model. *Lancet*. 2011;377(9761):219-227.

120. Magee LA, Helewa ME, Moutquin JM, et al. SOGC guidelines; diagnosis, evaluation and management of the hypertensive disorders of preganacy. *J Obstet Gynaecol Can*. 2008;30(3 Suppl):S1-S48.

121. von Dadelszen P, Payne B, Li J, et al. Prediction of adverse maternal outcomes in preeclampsia: Development and validation of the FullPIERS model EDITORIAL COMMENT. *Obstet Gynecol Surv*. 2011;66(5):267-268.

122. Payne B, Hodgson S, Hutcheon JA, et al. Performance of the fullPIERS model in predicting adverse maternal outcomes in pre-eclampsia using patient data from the PIERS (pre-eclampsia integrated estimate of RiSk) cohort, collected on admission. *BJOG: An International Journal of Obstetrics & Gynaecology*. 2013;120(1):113-118.

123. von Dadelszen P, Payne B, li J, et al. Predicting adverse maternal outcomes in pre-eclampsia: The fullPIERS model - development and validation. *Lancet*. 2011;377(9761):185.

124. Sibai BM. The HELLP syndrome (hemolysis, elevated liver enzymes, and low platelets): Much ado about nothing? *Am J Obstet Gynecol*. 1990;162(2):311-316.

125. Brown B, Cochran SW, Helmer O. An evaluation of methodology of delphi technique. *Biometrics*. 1967;23(3):600-&.

126. Steyerberg EW, Borsboom GJ, van Houwelingen HC, Eijkemans MJ, Habbema JD. Validation and updating of predictive logistic regression models: A study on sample size and shrinkage. *Stat Med*. 2004;23(16):2567-2586.

127. Toll DB, Janssen KJ, Vergouwe Y, Moons KG. Validation, updating and impact of clinical prediction rules: A review. *J Clin Epidemiol*. 2008;61(11):1085-1094.

128. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: A framework for traditional and novel measures. *Epidemiology*. 2010;21(1):128-138.

129. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982;143(1):29-36.

130. Deeks J, Altman D. Statistics notes - diagnostic tests 4: Likelihood ratios. *Br Med J*. 2004;329(7458):168-169.

131. Janes H, Pepe MS, Gu W. Assessing the value of risk predictions by using risk stratification tables. *Ann Intern Med*. 2008;149(10):751-760.

132. Dissanayake VH, Morgan L, Broughton Pipkin F, et al. The urine protein heat coagulation test--a useful screening test for proteinuria in pregnancy in developing countries: A method validation study. *BJOG: An International Journal of Obstetrics & Gynaecology*. 2004;111(5):491-494.

133. D'Agostino RB,Sr, Grundy S, Sullivan LM, Wilson P, CHD Risk Prediction G. Validation of the framingham coronary heart disease prediction scores: Results of a multiple ethnic groups investigation. *JAMA*. 2001;286(2):180-187.

134. Austin PC, Tu JV. Automated variable selection methods for logistic regression produced unstable models for predicting acute myocardial infarction mortality. *J Clin Epidemiol*. 2004;57(11):1138-1146.

135. Payne BA, Hutcheon JA, Ansermino JM, et al. A risk prediction model for the assessment and triage of women with hypertensive disorders of pregnancy in low-resourced settings: The miniPIERS (pre-eclampsia integrated estimate of RiSk) multi-country prospective cohort study. *PLoS Medicine / Public Library of Science*. 2014;11(1):e1001589.

136. Cook NR, Moons KG, Harrell FE,Jr. Assessing predictive performance beyond the framingham risk score. *JAMA*. 2010;303(14):1368-1369.

137. Steyerberg EW. Presentation FOrmats. In: Gail M, Krickberg K, Samet J, Tsiastis A, Wong W, eds. *Clinical prediction models: A practical approach to development, validating and updating.* 1st ed. New York: Springer; 2009:313-331.

138. Altman D. Problems in dichotomizing continuous-variables. *Am J Epidemiol*. 1994;139(4):442-442.

139. Altman D, Royston P. Statistics notes - the cost of dichotomising continuous variables. *Br Med J*. 2006;332(7549):1080-1080.

140. Shariat SF, Capitanio U, Jeldres C, Karakiewicz PI. Can nomograms be superior to other prediction tools?. *BJU Int*. 2009;103(4):492-495.

141. Costantine MM, Fox K, Byers BD, et al. Validation of the prediction model for success of vaginal birth after cesarean delivery. *Obstetrics & Gynecology*. 2009;114(5):1029-1033.

142. Dunsmuir D, Payne B, Cloete G, et al. Development of mHealth applications for pre-eclampsia triage. *IEEE Journal Biomed and Health Inform*. 2014;0(0):0.

143. Ryan M. *Using discrete choice experiments to value health and health care.* 1st ed. NY: Springer; 2008.

144. Community level interventions for pre-eclampsia (CLIP). https://pre-empt.cfri.ca/treatment/clip-trial. Updated 2014. Accessed January/01, 2014.

145. Hudson J, Nguku SM, Sleiman J, et al. Usability testing of a prototype phone oximeter with healthcare providers in high- and low-medical resource environments. *Anaesthesia*. 2012;67(9):957-967.

146. Muhlenbruch K, Heraclides A, Steyerberg EW, Joost HG, Boeing H, Schulze MB. Assessing improvement in disease prediction using net reclassification improvement: Impact of risk cut-offs and number of risk categories. *Eur J Epidemiol*. 2013;28(1):25-33.

147. Steyerberg EW, Pencina MJ, Lingsma HF, Kattan MW, Vickers AJ, Van Calster B. Assessing the incremental value of diagnostic and prognostic markers: A review and illustration. *Eur J Clin Invest*. 2012;42(2):216-228.

148. Pencina MJ, D'Agostino RB,Sr, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med*. 2011;30(1):11-21.

149. Pencina MJ, D'Agostino RB, Pencina KM, Janssens AC, Greenland P. Interpreting incremental value of markers added to risk prediction models. *Am J Epidemiol*. 2012;176(6):473-481.

150. Pencina MJ, D'Agostino RB, Vasan RS. Statistical methods for assessment of added usefulness of new biomarkers. *Clinical Chemistry & Laboratory Medicine*. 2010;48(12):1703-1711.

151. Kerr KF, Wang Z, Janes H, McClelland RL, Psaty BM, Pepe MS. Net reclassification indices for evaluating risk prediction instruments: A critical review. *Epidemiology*. 2014;25(1):114-121.

152. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation*. 2007;115(7):928-935.

153. Cook NR. Statistical evaluation of prognostic versus diagnostic models: Beyond the ROC curve. *Clin Chem*. 2008;54(1):17-23.

154. von Dadelszen P, Firoz T, Donnay F, et al. Preeclampsia in low and middle income countries-health services lessons learned from the PRE-EMPT (PRE-eclampsia-eclampsia monitoring, prevention and treatment) project. *Journal of Obstetrics & Gynaecology Canada: JOGC*. 2012;34(10):917-926.

155. Courvoisier DS, Combescure C, Agoritsas T, Gayet-Ageron A, Perneger TV. Performance of logistic regression modeling: Beyond the number of events per variable, the role of data structure. *J Clin Epidemiol*. 2011;64(9):993-1000.

156. Leushuis E, van der Steeg JW, Steures P, et al. Prediction models in reproductive medicine: A critical appraisal. *Hum Reprod Update*. 2009;15(5):537-552.

157. Coppus SF, van der Veen F, Opmeer BC, Mol BW, Bossuyt PM. Evaluating prediction models in reproductive medicine. *Human Reproduction*. 2009;24(8):1774-1778.

158. Steyerberg EW, Schemper M, Harrell FE. Logistic regression modeling and the number of events per variable: Selection bias dominates. *J Clin Epidemiol*. 2011;64(12):1464-1465.

159. Brunelli A, Rocco G. Internal validation of risk models in lung resection surgery: Bootstrap versus training-and-test sampling. *Journal of Thoracic & Cardiovascular Surgery*. 2006;131(6):1243-1247.

160. Smith GC, Seaman SR, Wood AM, Royston P, White IR. Correcting for optimistic prediction in small data sets. *Am J Epidemiol*. 2014;180(3):318-324.

161. Austin PC, Tu JV. Bootstrap methods for developing predictive models. *The American Statistician*. 2004;58(2):131-137.

162. Steyerberg EW, Eijkemans MJ, Van Houwelingen JC, Lee KL, Habbema JD. Prognostic models based on literature and individual patient data in logistic regression analysis. *Stat Med*. 2000;19(2):141-160.

163. Harrell FE,Jr, Shih YC. Using full probability models to compute probabilities of actual interest to decision makers. *Int J Technol Assess Health Care*. 2001;17(1):17-26.

Appendices

Appendix A miniPIERS data collection form

MATERNAL—CONFIRMATION OF ELIGIBILITY

Inclusion Criteria:

1. **PIERS Eligibility Criteria fulfilled upon and/or during any admission: (mark ALL that apply)**
   (see Working Protocol Pg. 5 for definitions)
   Pre-eclampsia:
   - ○ **Hypertension and proteinuria**
   - ○ **Hypertension and hyperuricaemia**
   - ○ **HELLP syndrome**
   - ○ **Superimposed Pre-eclampsia: pre-existing hypertension with new proteinuria**
   - ○ **Superimposed Pre-eclampsia: pre-existing hypertension with accelerated proteinuria (in the presence of chronic renal disease)**
   - ○ **Superimposed Pre-eclampsia: pre-existing hypertension with accelerated hypertension**
   - ○ **Superimposed Pre-eclampsia: pre-existing hypertension with new hyperuricaemia**

   Other Hypertensive Disorders of Pregnancy:
   - ○ **Pre-existing/chronic hypertension (diagnosed pre-pregnancy or < 20 weeks' gestation)**
   - ○ **Non-proteinuric gestational hypertension (diagnosed after 20 weeks' gestation)**
   - ○ **Partial HELLP syndrome (Examples: HEL, ELLP, or HLP)**

Exclusion Criteria:

Please also ensure that **NONE** of the following **EXCLUSION** criteria have been fulfilled:

- ○ **Occurrence of any element of the combined adverse maternal outcome** *(detailed on pg. 35 of Working Protocol)* **prior to collection of the predictor variables**

- ○ **Occurrence of any element of the combined adverse maternal outcome** *(detailed on pg. 35 of Working Protocol)* **prior to fulfillment of the eligibility criteria**

- ○ **Confirmed positive HIV/AIDS status with absolute CD4 count <250 or presence of AIDS defining illness**

- ○ **Admission to hospital in spontaneous labour**

If either of these exclusion criteria has occurred, this patient is **NOT ELIGIBLE** for inclusion in the PIERS project, and data need not be collected.

MATERNAL--DAY OF FIRST ADMISSION **SEGMENT 1**

Section A: Admission information

1. **During this pregnancy has the woman been previously admitted to this hospital for pre-eclampsia (suspected or confirmed)?**

○ Unknown

○ No → If **no**, please record the admission information for this **present** (and first) admission below. Then continue to **section B, question 2**.

○ Yes→ If **yes**, please record the **admission information** for **all previous** admissions to this hospital. Then continue to **section B, question 2**.

- Information for the **first** admission may be recorded below.
- Information for **all subsequent** re-admissions (including this present admission) may be recorded on separate "Maternal Re-admission" insert forms.
- Please ensure to also record any data available from **each and all** of these admissions in ***Tables 1-3*** (located in Segment 2 for the first admission, and in the "Maternal Re-admission" insert forms for all subsequent re-admissions), and in ***Segments 3***--Maternal Outcomes/Delivery ***and 4***--Neonatal Outcomes, where applicable.

○**First admission (admission #1)**

Date and time of admission **Date of discharge**

2 0 y y m m d d 2 4 : 0 0 → 2 0 y y m m d d

Year Month Day 24 hour clock Year Month Day

(a) Was the patient transferred from another institution? ○No ○Yes→ If **yes**, what was the **date and time** of **admission** at this other institution?

2 0 y y m m d d 2 4 : 0 0

Year Month Day 24 hour clock

*Please ensure to also record data available from this other institution in the following form segments, where applicable:

(**2**--Clinical Assessments and Labs, under "Other Site"; **3**--Maternal Outcomes/Delivery; **4**--Neonatal Outcomes.)

(b) Present Pregnancy Weight: ○ lb ○ Not specified

(c) At the time of this first admission, does ha ny of the foll g in this pregnancy?

	NOT SPECIFIED	NO	YES
i) Gestational diabetes?	○	○	○
ii) A history of smoking (any amount)?	○	○	○
iii) A history of oral tobacco use (Gutka)?	○	○	○
iv) A history of illicit drug use? (including: cocaine, heroin, marijuana, methamphetamines, ecstasy)	○	○	○

Section B: Identifiers

2. Mother's ethnicity *(choose the most appropriate **one**)*:

○Caucasian ○South Asian ○Pacific Islanders/Maori
○Black ○1st Nations ○Australian Aborigines ○ Not specified
○East Asian ○Latino ○Arabic/Middle Eastern
○Other_____
specify

3. Estimated date of delivery (EDD) by LMP or ultrasound: 2 0 y y m m d d

Year Month Day

4. Number of fetuses: ☐ current pregnancy; must be ≥ 1)

5. Pre-pregnancy weight: ☐☐☐.☐ kg **OR** ☐☐☐ lbs ○ Not specified

6. Height: ☐☐☐.☐ cm **OR** ☐☐☐ inches ○ Not specified

7. Gravidity ☐☐☐ **Preterm** ☐☐ **Abortion** ☐☐ **Living** ☐☐

☐☐ ☐☐ ☐☐ 34^0-36^6 weeks

☐☐ <34 weeks

8. Parity: ☐☐

(Number of previous deliveries of fetus at ≥ 20 weeks' gestation or ≥ 500 g birthweight)

Section C: Medical History

9. Does the woman's past obstetric and/or medical history include any of the following conditions, existing prior to or before 20 weeks' gestation in the current pregnancy?

(Please mark answer for each, or if none apply, mark "None") ○**NONE**

	NOT SPECIFIED		NO	YES
a) Gestational hypertension in a previous pregnancy*	○	○	○	○N/A--1st pregnancy
b) Gestational proteinuria in a previous pregnancy†	○	○	○	○N/A--1st pregnancy
c) Pre-gestational renal disease	○	○	○	
→If yes, **specify**:_____				
d) Pre-gestational hypertension	○	○	○	
e) Pre-gestational diabetes	○		○	○
e) Suspected or confirmed infection with malaria?				
Primary infection	○	○	○	
Recurrent infection	○		○	○

Table 1. Clinical Assessments (Part A):

Test date (yyyy/mm/dd)			**20**yy /mm /dd	**20**yy /mm /dd
Location	**Other site** (if transferred)	**Pre-admission** (within 12 hrs)	**Study site** First 24 hours	**Study site**
	O None	O None		
Clinical Assessments				
sBP (mmHg) BP measurement with highest sBP per date	sBP / dBP O NS	sBP / dBP O NS	sBP / dBP O NS	sBP / dBP O NS
dBP (mmHg) BP measurement with highest dBP per date	sBP / dBP O NS	sBP / dBP O NS	sBP / dBP O NS	sBP / dBP O NS
Temperature (°C) Highest reading per date	O NS	O NS	O NS	O NS
SaO_2 (%) lowest reading per date	O NS	O NS	O NS	O NS
Dipstick protein/Urine protein heat coagulation test highest reading per date	**Neg.** **Trace** ☐ + O NS	**Neg.** **Trace** ☐ + O NS	**Neg.** **Trace** ☐ + O NS	**Neg.** **Trace** ☐ + O NS
Symptoms				
Severe nausea	O **Yes** O **No**	O **Yes** O **No**	O **Yes** O **No**	O **Yes** O **No**
Severe vomiting	O **Yes** O **No**	O **Yes** O **No**	O **Yes** O **No**	O **Yes** O **No**
Headache	O **Yes** O **No**	O **Yes** O **No**	O **Yes** O **No**	O **Yes** O **No**
Visual disturbances	O **Yes** O **No**	O **Yes** O **No**	O **Yes** O **No**	O **Yes** O **No**
Right upper quadrant pain	O **Yes** O **No**	O **Yes** O **No**	O **Yes** O **No**	O **Yes** O **No**
Epigastric pain	O **Yes** O **No**	O **Yes** O **No**	O **Yes** O **No**	O **Yes** O **No**
Vaginal bleeding	O **Yes** O **No**	O **Yes** O **No**	O **Yes** O **No**	O **Yes** O **No**
Abdominal pain	O **Yes** O **No**	O **Yes** O **No**	O **Yes** O **No**	O **Yes** O **No**
Chest pain	O **Yes** O **No**	O **Yes** O **No**	O **Yes** O **No**	O **Yes** O **No**
Dyspnoea	O **Yes** O **No**	O **Yes** O **No**	O **Yes** O **No**	O **Yes** O **No**

* If response was positive for any of these events after admission, a "Maternal - Adverse Outcome Report" Insert Package **must** be completed.
NS = not specified / not available

Table 1. Clinical Assessments (Part B):

Test date (yyyy/mm/dd)				**20**yy /mm /dd			**20**yy /mm /dd		
Location	**Other site** (if transferred)			**Pre-admission** (within 12 hrs)			**Study site** First 24 hours			**Study site**		
	○ None			○ None								
Adverse Conditions:												
Units of blood or blood products transfused * (# of units)	○None ○Indicated but none received			○None ○Indicated but none received			○None ○Indicated but none received			○None ○Indicated but none received		
	RBCs			RBCs			RBCs			RBCs		
	Cryo			Cryo			Cryo			Cryo		
	FFP			FFP			FFP			FFP		
	Platelet			Platelet			Platelet			Platelet		
Eclamptic seizures (#)*	# of seizures			# of seizures			# of seizures			# of seizures		
Pulmonary oedema *	○**Yes** ○ **No** ○ Clinical dx only			○**Yes** ○ **No** ○ Clinical dx only			○**Yes** ○ **No** ○ Clinical dx only			○**Yes** ○ **No** ○ Clinical dx only		

Table 2. Maternal Lab Investigations (complete all available fields or select 'NS' to indicate a test was not specified/not available at the indicated time point):

Test date (yyyy/mm/dd)			**20**yy /mm /dd	**20**yy /mm /dd
Time **24:00 (24 hr. clock)**	:	:	24:00	24:00
Location	**Other site** (if transferred)	**Pre-admission** (within 2 weeks)	**Study site**	**Study site**
	O None	O None		
Hematological:				
White blood cells WBC **(10^9 cells/L)**	ONS	ONS	ONS	ONS
Platelets (10^9/L)	ONS	ONS	ONS	ONS
Mean platelet volume MPV **(fL)**	ONS	ONS	ONS	ONS
Hematocrit (express % as decimal)	ONS	ONS	ONS	ONS
Coagulation:				
INR	ONS	ONS	ONS	ONS
Fibrinogen (g/L)	ONS	ONS	ONS	ONS
aPTT (sec)	ONS	ONS	ONS	ONS
Chemistry:				
Serum creatinine (µM)	ONS	ONS	ONS	ONS
Uric acid (µM)	ONS	ONS	ONS	ONS
AST or ASAT (U/L)	ONS	ONS	ONS	ONS
ALT or ALAT (U/L)	ONS	ONS	ONS	ONS
LDH (U/L)	ONS	ONS	ONS	ONS
Serum albumin (g/L)	ONS	ONS	ONS	ONS

Unconjugated bilirubin (µM)	○NS	○NS	○NS	○NS
Conjugated bilirubin (µM)	○NS	○NS	○NS	○NS
Total bilirubin (µM)	○NS	○NS	○NS	○NS
Random glucose (mM)	○NS	○NS	○NS	○NS

NS = not specified / not available

Table 2 (continued). Maternal Lab Investigations (complete all available fields or select 'NS' to indicate a test was not specified/not available at the indicated time point):

○ **None performed:** No urine tests (Random or 24 hour) were performed during this **entire** admission to the study site.

Test date(yyyy/mm/dd)			**20**yy /mm /dd	**20**yy /mm /dd
Location	**Other site** (if transferred)	**Pre-admission** (within 2 weeks)	**Study site**	**Study site**
	○ None	○ None		
Random urinalysis:				
Was a random urinalysis performed?	○**Yes** ○ **No**	○**Yes** ○ **No**	○**Yes** ○ **No**	○**Yes** ○ **No**
Time 24:00 (24 hr. clock)	:	:	24:00	24:00
Random urinary creatinine (mmol/L)	○NS	○NS	○NS	○NS
Random urinary protein (mg/L)	○NS	○NS	○NS	○NS
Protein:creatinine ratio (mg/mmol)	○NS	○NS	○NS	○NS
Random urinary albumin (mg/L)	○NS	○NS	○NS	○NS
Albumin:creatinine ratio (mg/mmol)	○NS	○NS	○NS	○NS
Date results reported	**20**yy /mm /dd	**20**yy /mm /dd	**20**yy /mm /dd	**20**yy /mm /dd
24 hour urinalysis				
Was a 24 hour urinalysis performed?	○**Yes** ○ **No**	○**Yes** ○ **No**	○**Yes** ○ **No**	○**Yes** ○ **No**
Time collection started	:	:	24:00	24:00

24:00 (24 hr. clock)				
Volume (L/d)	○NS	○NS	○NS	○NS
[Conc.] of urinary creatinine (mmol/L)	○NS	○NS	○NS	○NS
[Conc.] of urinary protein (g/L)	○NS	○NS	○NS	○NS
Urinary creatinine excretion (mmol/d)	○NS	○NS	○NS	○NS
Urinary protein excretion (g/d)	○NS	○NS	○NS	○NS
Date results reported	**20**yy /mm /dd	**20**yy /mm /dd	**20**yy /mm /dd	**20**yy /mm /dd

NS = not specified / not available

Table 3. Fetal Assessments (complete all available fields or select 'NS' to indicate a test was not specified/not available at the indicated time point):

Form for: ○ **Singleton or Baby A**

○ **None performed:** No fetal assessments (ultrasound or fetal heart rate trace) were performed during this **entire** admission to the study site.

Test date (yyyy/mm/dd)			**20**yy /mm /dd
Location	**Other site** (if transferred)	**Pre-admission** (within 2 weeks)	**Study site**
	○ None	○ None	
Ultrasound:			
Was an ultrasound performed?	○**Yes** ○ **No**	○**Yes** ○ **No**	○**Yes** ○ **No**
Time 24:00 (24 hr. clock)	:	:	24 : 00
Amniotic fluid index (mm) for single fetus	mm	mm	mm
	○NS	○NS	○NS
Deepest amniotic fluid pocket (mm) for multiple fetuses	mm	mm	mm
	○NS	○NS	○NS
Estimated fetal weight (grams)	g	g	g
	○NS	○NS	○NS
Abdominal circumference (mm)	mm	mm	mm

	○NS	○NS	○NS	
Umbilical artery Doppler Diastolic flow	○ Present ○ Absent ○ Reversed ○NS	○ Present ○ Absent ○ Reversed ○NS	○ Present ○ Absent ○ Reversed ○NS	
Fetal heart rate trace:				
Was a CTG or NST performed?	○**Yes** ○ **No**	○**Yes** ○ **No**	○**Yes** ○ **No**	
Time 24:00 (24 hr. clock)	:	:	24 : 00	24 : 00
Fetal heart rate trace Rating visual	○Normal ○Suspicious/ Atypical ○Pathological /Abnormal ○NS	○Normal ○Suspicious/ Atypical ○Pathological /Abnormal ○NS	○Normal ○Suspicious/ Atypical ○Pathological /Abnormal ○NS	○Normal ○Suspicious/ Atypical ○Pathological /Abnormal ○NS

NS = not specified / not available

MATERNAL OUTCOMES / DELIVERY **SEGMENT 3**

Section A: Events during course of pregnancy

1. **During this pregnancy, is the woman receiving (or has she received) any of the following?**

a) Corticosteroids at any time during pregnancy (Betamethasone or Dexamethasone):

○Unknown ○No ○Yes → If yes, **specify date** of **first** dose:

Year **Month** **Day**

2 0 y y m m d d

b) Antihypertensives during any admission: (see Appendix A of Working Protocol for a complete list of drugs to collect)

○Unknown ○No ○Yes → If yes, **specify antihypertensive agents used**:

(*select **ALL** that apply*)

○**Oral** Labetalol
○**Oral** Methyldopa
○**Oral** hydralazine
○Nifedipine long acting (XL or Oros)
○Nifedipine intermediate acting (PA or Retard)
○Parenteral labetalol (intravenous)
○Parenteral hydralazine (intravenous/intramuscular)
○Nifedipine capsule
○Prazosin
○Other(s)_____

specify

_____

c) $MgSO_4$ during <u>any admission</u>:

○Unknown ○No ○Yes

d) Other(s) during <u>any admission</u>: (see <u>Appendix A</u> of Working Protocol for a complete list of drugs to collect)

○Unknown ○No ○Yes → If yes, **specify agents used**:

*<u>(select **ALL** that apply)</u>*

○Heparin → If yes, was heparin used for post-surgical prophylaxis? ○No ○Yes
○Warfarin/Coumadin®
○Insulin
○Anticonvulsant(s)
○Antiretroviral(s)
○Antimalarial(s)
○Oral hypoglycaemics
○Calcium supplementation
○Other(s)_____

specify

_____

e) Medications/drugs received as part of participation in other studies/trials (e.g. CHIPS, TIPPS, etc.)

○Unknown ○No ○Yes → If yes, **specify** the **study/trial,** the **arm/treatment group** the patient is enrolled in:

○_____ _____
study/trial name **arm/treatment group**

○_____ _____
study/trial name **arm/treatment group**

<u>Section B: Adverse maternal outcomes</u>

2. Did any of the following <u>serious</u> maternal complications develop between first admission and ultimate hospital discharge?

a) Mortality (maternal) **NO** **YES**
○ ○ if **yes**→complete an "Adverse Outcome Report" insert package

<u>Hepatic:</u>

b) Hepatic dysfunction *(INR greater than 1.2 in the **absence** of disseminated intravascular coagulation (DIC)* <u>or</u> treatment with Warfarin (Coumadin®).*

If DIC **is** present, <u>or</u> if the patient is receiving Warfarin (Coumadin®), then hepatic dysfunction is defined as mixed hyperbilirubinaemia (> 17 µM), **or** hypoglycaemia (< 2.5 mM) in the absence of insulin.*

**DIC is defined on pg. 36 of the Working Protocol)*

NO **YES**
○ ○ if **yes**→complete an "Adverse Outcome Report" insert package

c) Hepatic haematoma/rupture

(Defined by the presence of blood collection under the hepatic capsule as confirmed by ultrasound or at laparotomy.)

NO **YES**
○ ○ if **yes**→complete an "Adverse Outcome Report" insert package

<u>Central Nervous System:</u>

d) Glasgow coma score < 13 *(GCS is a scale that assesses the degree of coma in patients with craniocerebral injuries and also assesses brain function, brain damage, and patient progress. It is a composite measure that combines separate scores for the patient's eye opening, verbal, and motor responses. Please refer to the GCS scoring system located on pg. 36 of the Working Protocol for details.)*

NO **YES**

○ ○ if **yes**→complete an "Adverse Outcome Report" insert package

e) Stroke *(Acute neurological event with deficits lasting greater than 48 hours)*

NO **YES**

○ ○ if **yes**→complete an "Adverse Outcome Report" insert package

f) Cortical blindness (*Loss of visual acuity in the presence of intact pupillary response to light)*

NO **YES**

○ ○ if **yes**→complete an "Adverse Outcome Report" insert package

g) Other severe neurological events ***(Examples:*** *Reversible Ischaemic Neurolic Deficit (RIND), retinal detachment. Please refer to pg. 37 of the Working Protocol for definitions)*

NO **YES**

○ ○ if **yes**→ **specify each event below**

○_____ → complete an "Adverse Outcome Report" insert package **neurological event**

○_____ → complete an "Adverse Outcome Report" insert package **neurological event**

CNS investigations: *(applies to all CNS outcomes listed in items 'd' through 'g' above)*

Please indicate **investigation(s) conducted** and **record any findings** below:

Investigation:	**Findings:**
○ MR (magnetic resonance)	_____
○ CT (computerized tomography)	_____
	_____

Renal:

h) Acute renal insufficiency: a) in absence of underlying renal disease *(creatinine >150 μM) or b) regardless of renal disease creatinine >200 uM:*

NO **YES**

○ ○ if **yes**→complete an "Adverse Outcome Report" insert package

i) Dialysis *(May include haemodialysis or peritoneal dialysis.)*

NO **YES**

○ ○ if **yes**→complete an "Adverse Outcome Report" insert package

Heamatological:

j) Postpartum Hemorrhage requiring transfusion or hysterectomy

NO **YES**

○ ○ if **yes**→complete an "Adverse Outcome Report" insert package

k) platelet count <50,000 without transfusion

NO **YES**

○ ○if **yes**→complete an "Adverse Outcome Report" insert package

Cardiovascular:

l) Positive inotropic support required *(The use of vasopressors to maintain a systolic blood pressure of >90 mmHg or a mean arterial pressure >70mmHg.)*

NO **YES**

○ ○ if **yes**→complete an "Adverse Outcome Report" insert package

m) Infusion of a third injectable antihypertensive (nitroprusside, nitroglycerine/glyceryl trinitrate (NTG/GTN), diazoxide, and/or prazosin)

(Indication that patient has received infusion of a 3rd injectable antihypertensive because of uncontrollable hypertension.)

NO **YES**

○ ○ if **yes**→complete an "Adverse Outcome Report" insert package

n) severe uncontrolled hypertension *(requirement of 3 or more different antihypertensives administered by any route within 12 hours)*

NO **YES**

○ ○if **yes**→complete an "Adverse Outcome Report" insert package

o) Myocardial ischaemia/infarction

*(**Criteria for myocardial ischaemia:** ECG changes (ST segment elevation or depression) without enzyme changes.*

***Criteria for established myocardial infarction:** Any one of the following criteria satisfies the diagnosis for established MI: 1) Development of new pathologic Q waves on serial ECGs. The patient may or may not remember previous symptoms. Biochemical markers of myocardial necrosis may have normalized, depending on the length of time that has passed since the infarct developed. 2) Pathological findings of a healed or healing MI.*

***Criteria for acute, evolving or recent myocardial infarction:** Either one of the following criteria satisfies the diagnosis for an acute, evolving or recent MI: 1) Typical rise and gradual fall (troponin) or more rapid rise and fall (CK-MB) of biochemical markers of myocardial necrosis with at least one of the following: a) ischaemic symptoms; b) development of pathologic Q waves on the ECG; c) ECG changes indicative of ischaemia (ST segment elevation or depression); or d) coronary artery intervention (e.g., coronary angioplasty). 2) Pathological findings of an acute MI.)*

NO **YES**

○ ○ if **yes**→complete an "Adverse Outcome Report" insert package

Respiratory:

p) Require $\geq$ 50% O_2 for > 1hour

NO **YES**

○ ○ if **yes**→complete an "Adverse Outcome Report" insert package

q) Intubation (vent, EIT, CPAP) other than for Caesarean section

NO **YES**

○ ○ if **yes**→complete an "Adverse Outcome Report" insert package

r) Severe breathing difficulty including chest pain or dyspnoea, crackles in the lungs or SaO2 <90% *(suspected but not confirmed pulmonary oedema)*

NO **YES**

○ ○ if **yes**→complete an "Adverse Outcome Report" insert package

Other adverse maternal events:

NO **YES**

○ ○ if **yes**→**specify each event below**

○_____ → complete an "Adverse Outcome Report" insert package

adverse event

○_____ → complete an "Adverse Outcome Report" insert package

adverse event

Section C: Delivery

3. Please indicate primary reasons for delivery by marking the appropriate boxes:

*(please mark **ALL** that apply):*

○ Maternal condition

○ Fetal condition

○ Booked elective caesarean section

○ Spontaneous labour → if **yes** please specify date and time of onset of first stage of labour: | | | | | | 2 0 | y | y | m | m | d | d | | | : | | | |

○ Unspecified

4. Was labour induced?

○**No** ○**Yes**

5. What type of analgesia or anaesthesia was administered? *(Please mark **ALL** that apply)*

○None
○Narcotics (iv/im)
○General anaesthesia
○Regional anaesthesia (epidural, spinal, combined spinal epidural (CSE))
○Nitrous oxide (N_20_2)/Entonox

6. Was there evidence of placental abruption?

○**No** ○**Yes** If **yes**→ **specify** the primary manner by which placental abruption was diagnosed:

○**Clinical**
○**Ultrasound confirmed**
○**Delivery or placental pathology confirmed**

→ complete an "Adverse Outcome Report" insert package

7. For day of delivery *(i.e. between 12 a.m. (24:00) and 11:59 p.m. (23:59) on the day of delivery)*, record:

a) the blood pressure measurements containing the highest systolic blood pressures recorded before (antepartum, including active labour) and after (postpartum) parturition

AND

b) the blood pressure measurements containing the highest diastolic blood pressures recorded before (antepartum, including active labour) and after (postpartum) parturition.

		Antepartum	Postpartum
a)	**sBP (mmHg)** BP measurement with highest sBP on day of delivery	sBP / dBP ○NS	sBP / dBP ○NS
b)	**dBP (mmHg)** BP measurement with highest dBP on day of delivery	sBP / dBP	sBP / dBP ○NS

Section D: Events after delivery

8. To where was the woman discharged?

a) Home ○No ○Yes→ If yes, indicate **date of ultimate discharge home after delivery** (**or date of death** if maternal death occurred prior to discharge)**:**

2 0 y y m m d d
Year Month Day

b) Another institution ○No ○Yes → If yes, indicate the **name of the institution to which she was transferred** and **date of transfer:**

_____
Name of institution

2 0 y y m m d d
Year Month Day

→ If yes, indicate the date of the woman's **final discharge HOME** from this other institution (**or date of death** if maternal death occurred prior to discharge)**:**

2 0 y y m m d d ○Not known
Year Month Day

Please REVISIT the Eligibility Criteria listed on Page 1 of this Data Collection Form. Ensure ALL Eligibility Criteria fulfilled at any point during the course of the woman's

NEONATAL OUTCOMES **SEGMENT 4**

1. What was the mode of delivery? (*please choose **one***)

○Spontaneous vaginal ○Operative vaginal ○Caesarean section→ ○laboured ○non-laboured

2. Baby's date and time of birth: 2 0 y y m m d d Year 2 4 : 0 0 Month Day 24 hour clock

3. Gender: ○ Male ○ Female ○Indeterminate

4. Status of the baby at birth:

○Alive ○Stillborn→ if stillborn, **what was the primary cause of death?** *(mark **ALL** that apply)*

a. Congenital anomaly ○ No ○ Yes → if yes, **specify**_____

b. Other ○ No ○ Yes → if yes, **specify**_____

c. Undetermined ○ No ○ Yes

5. Apgar score: at 5 minutes ○ Not recorded **at 10 minutes** ○ Not recorded

6. Birthweight: grams ○ Not recorded

If baby stillborn, then data collection is now complete. Thank you. Otherwise, please proceed.

7. Was cord taken for blood gases?

○Not recorded ○ No ○ Yes → if **yes**, indicate vessel blood sample(s) taken from and any results:

8. Did the baby have respiratory distress after the initial resuscitation/stabilisation?

○ No ○ Yes → if yes, **what was the cause?** *(Please mark **ALL** that apply)*

respiratory distress syndrome	○ No	○ Yes
meconium aspiration syndrome	○ No	○ Yes
pneumonia	○ No	○ Yes
pneumothorax/pneumomediastinum	○ No	○ Yes
transient tachypnoea of the newborn	○ No	○ Yes
other_____	○ No	○ Yes

specify

9. Did the baby receive supplemental oxygen after the initial resuscitation/stabilisation?

○ No ○ Yes → if yes, indicate **date and time oxygen first started:**

2 0 y y m m d d 2 4 : 0 0 year month day 24 hour clock

→ if yes, indicate **date and time oxygen finally stopped, or the date of discharge if sent home on oxygen:**

2 0 y y m m d d 2 4 : 0 0 year month day 24 hour clock

10. Did the baby receive intubation and ventilation via endotracheal tube after the initial resuscitation/stabilisation?

○ No ○ Yes → if yes, indicate **date and time endotracheal ventilation first started:**

2 0 y y m m d d 2 4 : 0 0 year month day 24 hour clock

→ if yes, indicate **date and time endotracheal ventilation finally stopped:**

2 0 y y m m d d 2 4 : 0 0 year month day 24 hour clock

11. Did the baby receive any other ventilatory support without intubation after the initial

resuscitation/stabilisation? *(nasopharyngeal, nasal cannula, mask, CPAP)*

○ No ○ Yes → if yes, indicate **date and time other support <u>first started</u>:**

2 0 y y m m d d 2 4 0 0
year month day 24 hour clock

→ if yes, indicate **date and time other support <u>finally stopped</u>:**

2 0 y y m m d d 2 4 0 0
year month day 24 hour clock

12. Did the baby receive surfactant?

○ No ○ Yes → if yes, indicate **date and time of <u>first dose</u>:**

2 0 y y m m d d 2 4 0 0
year month day 24 hour clock

→ if yes, **was the surfactant given for prophylaxis?** ○No ○Yes

Appendix B Components of the PIERS composite adverse maternal outcome with definition

Outcome	Definition
Mortality	Maternal death occurring within six weeks of pregnancy or if later, attributable to complications of pre-eclampsia
Hepatic dysfunction	INR >1.2 in the absence if DIC or treatment of Warfarin (DIC is defined as having both: abnormal bleeding and consumptive coagulopathy (i.e., low platelets, abnormal peripheral blood film, or one or more of the following: increased INR, increased PTT, low fibrinogen, of increased fibrin degradation products that are outside normal non-pregnancy ranges)
Hepatic hematoma or rupture	Blood collection under the hepatic capsule as confirmed by ultrasound or laparotomy
Glasgow coma score < 13	Based on GCS scoring system: Teasdale G, Jennet B. Assessment of coma and impaired consciousness: a practical scale. *Lancet* 1974; **2**:81-83
Stroke	Acute neurological event with deficits lasting longer than 48 hours
Cortical Blindness	Loss of visual acuity in the presence of intact papillary response to light
Rversible Ischaemic Neurologic Deficit (RIND)	Cerebral ischaemia lasting longer than 24 hrs but less than 48 hours revealed through clinical examination
Retinal detatchment	Separation of the inner layers of the retina from the underlying retinal pigment epithelium (RPE, choroid) and is diagnosed by opthamological exam
Acute renal insufficiency	For women with an underlying history of renal disease: defined as creatinine >200 uM; for patients with no underlying renal disease: defined as creatinine >150 uM
Dialysis	Including haemodialysis and peritoneal dialysis
Postpartum hemorrhage requiring transfusion or hysterectomy	Occurrence of PPH that required transfusion or hysterectomy
Placental Abruption	Any occurrence of abruption diagnosed

Outcome	Definition
	clinically or based on placental pathology report
Platelet count < 50,000 without blood transfusion	Measurement of platelet count recorded as less than 50,000 without patient being given a blood transfusion
Transfusion of blood products	Includes transfusion of any units of blood products: fresh frozen plasma (FFP), platelets, red blood cells (RBCs), cryoprecipitate (cryo) or whole blood
Positive inotropic support	The use of vasopressors to maintain a sBP > 90 mmHg or Mean Arterial pressure > 70 mmHg
Myocardial ischaemia/infarction	ECG changes (ST segment elevation or depression) without enzyme changes AND/OR any one of the following: 1)Development of new pathologic Q waves on serial ECGs. The patient may or may not remember previous symptoms. Biochemical markers of myocardial necrosis may have normalized, depending on the length of time that has passed since the infarct developed. 2) Pathological findings of an acute, healed or healing MI 3) Typical rise and gradual fall (troponin) or more rapid rise and fall (CK-MB) of biochemical markers of myocardial necrosis with at least one of the following: a) ischaemic symptoms; b) development of pathologic Q waves on the ECG; c) ECG changes indicative of ischaemia (ST segment elevation or depression); or d) coronary artery intervention (e.g., coronary angioplasty)
Eclampsia	Any episode of seizure antepartum, intrapartum or before postpartum discharge as follow-up beyond discharge is not possible
Require >50% oxygen for greater than one hour	Oxygen given at greater than 50% concentration based on local criteria for longer than 1 hour

Outcome	Definition
Intubation other than for Cesarean section	Intubation may be by ventilation, EIT or CPAP
Severe breathing difficulty	Suspected pulmonary oedema where x-ray confirmation unavailable may be diagnosed by presence of chest pain or dyspnoea, crackles in the lungs and SaO_2 <90%
Pulmonary Oedema	Clinical diagnosis with x-ray confirmation or requirement of diuretic treatment and SaO_2 <95%

Appendix C Discrete choice survey

This simple survey is designed to assess how people rank sensitivity vs. specificity and various clinical considerations in the context of caring for women with pre-eclampsia. To complete the survey mark an X beside the preferred option of each pair in the list provided. Please feel free to comment on your assumptions and considerations for each selection in the final column.

	A		B		Comments
1	Transfer of a women with mild pre-eclampsia for urgent care because she was misclassified as high-risk of outcomes	X	Non-urgent transfer of a women for care (taking >24hrs) who is at high risk of outcomes but was misclassified as low-risk		
2	Treating a woman with intramuscular $MgSO_4$ in the community who is low-risk of outcomes but was misclassified as high-risk, and then transferring her to care.		Transferring a woman to tertiary centre for care because she was misclassified as low-risk, then treating her with $MgSO_4$ on admission because she is actually high-risk.	X	
3	Treating a woman with intramuscular $MgSO_4$ in the community who is misclassified as high-risk of outcomes, and then transferring her to care		Providing counseling to a women who is misclassified as low-risk to rest and relax, then come back in a week for reassessment		
4	Providing counseling to a woman with blood pressure of 143/90 mmHg to rest and relax, then asking her to come back for reassessment in 48 hrs at the community clinic		Providing a dose of oral antihypertensive to a woman with blood pressure of 143/90 mmHg in the community, then asking her to transfer to care at hospital within 24hrs		
5	Providing counseling to a woman with blood pressure of 150/95 mmHg to rest and relax, then asking her to come back for reassessment in 48 hrs at the community clinic		Providing a dose of oral antihypertensive to a woman with blood pressure of 150/95 mmHg in the community, then asking her to transfer to care at hospital within 24hrs		
6	Non-urgent transfer of a women at 31 weeks gestation (taking >24hrs) who is misclassified as low-risk		Non-urgent transfer of a women at 36 weeks gestation (taking >24hrs) who is misclassified as low-risk		
7	Treatment of a women at 29 weeks gestation with $MgSO_4$ and an oral antihypertensive in the community who was misclassified as high-risk		Treatment of a women at 34 weeks gestation with $MgSO_4$ and an oral antihypertensive in the community who was misclassified as high-risk		

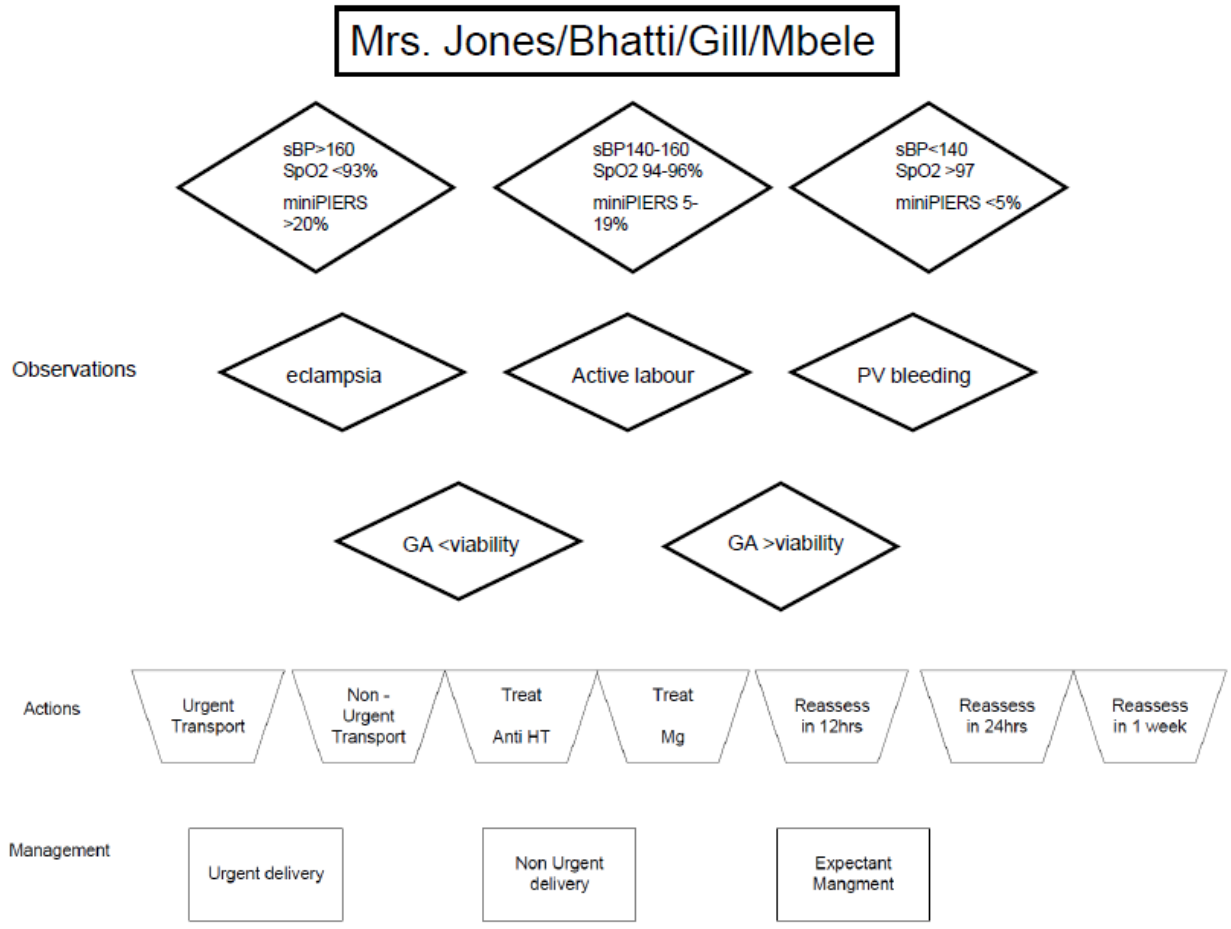
Appendix D Decision tree framework

Instructions

•Imagine you are a community or primary health centre worker. A woman comes to see you who is pregnant and you note she has high blood pressure. Map out, what observations or measurements you would want to take of this women, then what thresholds you would use to decide when to use the listed actions. Feel free to add additional actions or management options, if you feel something is missing.

•Example Observations/measurements: blood pressure, proteinuria, symptoms, miniPIERS score, gestational age, seizures, etc.

•Use hierarchy to show weighting of importance of various measurements and thresholds on the decision.



Appendix E Treatment and Referral trigger questionnaire

Indication	Definition	Should this be used in CLIP to necessitate transport?	Should this be used in CLIP to necessitate treatment?
1. Eclampsia	Any episode of seizure during the current pregnancy, or if women is postpartum, any episode of seizure in the postpartum period that has not been followed by treatment with $MgSO_4$	YES NO	With MgSO4?
2. Signs or symptoms of recent stroke	Recent onset of weakness on one side of the body or transient blindness	YES NO	With MgSO4? With antihypertensive?
3. Significant vaginal bleeding	New onset of vaginal bleeding with volume greater than a large spoonful	YES NO	With MgSO4?
4. Unconscious or unresponsive	For greater than 12 hours	YES NO	With MgSO4? With antihypertensive?
5. miniPIERS probability ≥ 25%	Based on miniPIERS predicted probability associated with significantly increased likelihood of adverse event within 48hrs (+LR>10)	YES NO	With MgSO4? With antihypertensive?
6. sBP ≥ 160 mmHg	Measured while woman is in supine position and confirmed with repeated measurement after 5 minutes.	YES NO	With MgSO4? With antihypertensive?

Additional questions regarding the decision algorithm:

1. Should we add a consideration for normotensive eclampsia as a trigger on the left-hand side of the diagram? How common is eclampsia without hypertension?
2. Should transport be stratified between basic and comprehensive (EmOC) facilities for mild/moderate and severe pre-eclampsia, respectively. This is currently described in the decision tree as a method of separating urgent and non-urgent referral.
3. Do we need to add in consideration of gestational age as part of the decision to recommend basic vs. EmOC facility transport?

Appendix F Research Midwife survey tool for assessment of POM application feasibility and acceptability at Tygerberg Hospital, Cape Town, South Africa

Post-study Questionnaire

These questions are designed to gain a better understanding of the acceptability of the mobile health tool for assessment of pregnant women in antenatal care at Tygerberg Hospital, Cape Town, South Africa. You have been invited to complete this survey because your input is essential. Please provide as much information as possible so that we may learn from your experience assessing women with the PIERS on the Move tool over the past year. Excerpts from your response may be used to inform improvements to the app or be used in publications to describe the potential for use of the app in a clinical setting.

All information recorded will be kept confidential and you will not be identified by name. You may chose not to respond at any time.

This questionnaire is expected to take roughly 30 minutes.

Instructions: Please respond to all questions to the best of your ability. If you are not sure about your response please mark so below the question.

1. How easy was the application to use?
   very easy ____
   somewhat easy ___
   neutral ____
   difficult ___
   very difficult ___
   unsure ____
   COMMENT:

2. What part(s) of the application was easiest to use?
3. What part(s) of the application was most challenging to use?
4. What strategies did you use to overcome these challenges?
5. How easy was the application to learn?
   very easy ____
   somewhat easy ___
   neutral ____
   difficult ___
   very difficult ___
   unsure ____
   COMMENT:

6. If you were to design a training program for other nurses/midwives for this application, what would be the main elements included?
7. Did any patients express positive views of the application, if so, what were those views?
8. Did any patients express negative views of the application, if so, what were those views?
9. Did patients ask you questions about this application? If so, what did they ask you?
10. What did you say to patients to explain the purpose of the application?
11. Do you think it would be possible to integrate use of a tool such as this application into clinical care in a hospital setting?
12. Would there be challenges to introducing this tool into routine clinical care? If so, what would these challenges be?
13. Who would benefit most from this application?
14. Do you have any other comments you feel are important to share with the research team?

Thank for your participation