Retrospective Validation of the PREMature Infant Index (PREMII) using a Real-World Database

Study protocol

Final version 1.3

June 26, 2018

PROTOCOL

TITLE:	***Retrospective Validation of the PREMature Infant Index (PREMII) using a Real-World Database***
PROTOCOL VERSION AND DATE:	Version 1.3, 26 June 2018
REGISTRATION NO.:	The study will not be registered on any study registries (i.e. ClinicalTrials.gov, ENCePP) given the purpose and design of the study.
SPONSOR/MAH:	Shire Development LLC and International Affiliates 725 Chesterbrook Boulevard, Wayne, PA 19087 USA
RESEARCH QUESTION:	Shire has developed the PREMature Infant Index (PREMII), a clinician-reported outcome instrument to assess an infant's overall functional status – defined as what an infant can do with respective to key functional areas as a reflection of the infant's overall health and maturation during a stay in the neonatal unit (NNU). Shire is planning to use this tool in an upcoming phase IIb/III clinical trial as a secondary clinical endpoint. Therefore the current study is being conducted to validate this tool against a real-world database.
OBJECTIVES:	The primary objective of this study is to evaluate the predictive validity of the PREMII score measured at 36 weeks of PMA with subsequent LOS in the NICU. The secondary objective is to assess the discriminate validity of the PREMII score measured at 36 weeks of postmenstrual age (PMA) with discharge from the NNU during that same week.
COUNTRY(-IES) OF STUDY:	The study will be conducted in Canada; however the data source for this study is from the United Kingdom.
AUTHOR:	Sujata Sarda Email: ssarda0@shire.com
CONTACT PERSON:	Sujata Sarda Director, Outcomes Research and Epidemiology Shire 45 Hayden Avenue, Lexington, MA 02421 +1 (781) 266-3440 ssarda0@shire.com

PROTOCOL SIGNATURE PAGE

Shire Approval

Signature: Sujata	**Date:** 26-June-2018
Sujata Sarda, Director ORE	*{Note: Signature date must not precede the approval date}*

Investigator/Vendor Name and Address: (please hand print or type)	*{The Investigator/Vendor completes the bottom section of the protocol signature page}* Michael N. Cook Ph.D. Principal, Epidemiology ICON plc

Signature: Michael Cook **Date:** 26-June-2018

1. TABLE OF CONTENTS

LIST OF APPENDICES

2. ABBREVIATIONS

LOS	length of stay
NNU	Neonatal unit
NNRD	National Neonatal Research Database
PMA	postmenstrual age
PREMII	PREMature Infant Index

3. RESPONSIBLE PARTIES

Protocol Author(s):

Sisi Wang, MPH
Epidemiologist
ICON plc.

Sheena Kayaniyil, PhD
Senior Epidemiologist
ICON plc.

Michael Cook, PhD
Principal, Health Economics and Epidemiology
ICON plc.

Sujata Sarda, PhD
Director, Outcomes Research and Epidemiology
Shire

Third Party Agency/Representative:

Professor Neena Modi, MD
Professor of Neonatal Medicine
Imperial College London

Michael Cook, PhD
Principal, Health Economics and Epidemiology
ICON plc.

The name and address of each third party vendor used in this study will be maintained in the investigator's and sponsor's files.

4. ABSTRACT

Title of the study: Retrospective Validation of the PREMature Infant Index (PREMII) using a Real-World Database

V1.3, 26 June 2018

Sujata Sarda, PhD

Director, Outcomes Research and Epidemiology

Shire

Rationale and background:

Reducing length of stay (LOS) in the neonatal unit (NNU) may be considered a clinically meaningful outcome for extremely premature infants. LOS is a function of whether or not an infant's overall health and functional status have improved to the degree to allow discharge from the NNU in the opinion of the infant's treating paediatrician. As such, variation in parents' ability to care for the infant at home coupled with variation in physician and hospital practice patterns across NNU may affect LOS for premature infants who are at the same level of functional status. As an alternative to LOS, Shire has developed the PREMature Infant Index (PREMII), a clinician-reported outcome to measure an infant's overall functional status on each day of a stay in the NNU. As the PREMII is intended to be used as a secondary clinical endpoint in an upcoming phase IIb/III clinical trial, it is important to generate evidence to support the validity of this newly-developed instrument in a real-world practice setting.

Research question and objectives:

The overarching aim of the study is to validate the PREMII in a representative sample of premature infants in a real-world practice setting.

- The primary objective of this study is to evaluate the predictive validity of the PREMII score measured at 36 weeks of PMA with subsequent LOS in the NICU.
- .
- The secondary objective is to assess the discriminate validity of the modified PREMI score measured at 36 weeks of postmenstrual age (PMA) with discharge from the NNU during that same week.

The exploratory objectives are as follows:

- Describe the distribution of infants at each level of each factor in the PREMII score by week of PMA.
- Evaluate the discriminate validity of the PREMII score measured at 40 weeks of PMA with discharge from the NICU (yes/no) during that same week.

Study design:

A retrospective database analysis will be conducted. Infants identified for inclusion from a database will be retrospectively followed from birth to discharge from the NNU (retrospective observational cohort study design).

Data Sources:

The United Kingdom (UK)'s National Neonatal Research Database (NNRD) will be used to obtain relevant data on infants during their stay in the NNU. This database was selected as a result of a feasibility assessment which examined potential real-world databases that could be used as a validation data source for the PREMII.

Following research ethics approval and confirmation of agreement to participate from all NNU leads, it is anticipated that the de-identified data from the NNRD will be received. All data will then be reviewed and cleaned as necessary.

Population:

The target population of this study will be extremely premature infants (<28 weeks of gestational age) admitted to a NNU after birth.

The source population will be all infants meeting the above criteria and having data on or after 01 January 2012 in the NNRD. It is expected that all infants will be from England..

All eligible infants need to have complete data on demographics (i.e. month of birth, year of birth, birth weight, gestational age and sex) and admission and discharge details (derived from minutes from birth)

Infants with congenital heart disease, congenital diaphragmatic hernia, or congenital malformations at admission, day of care, and at discharge will be excluded from the study. Congenital heart disease for this study includes having bulbus cordis anomalies of cardiac septal closure (excluding patent foramen ovale), or other congenital anomalies of heart or of circulatory system (excluding patent ductus arteriosus). Congenital malformation for this study includes having congenital musculoskeletal deformities of chest, congenital malformations of lung or congenital malformations of upper airway

No additional inclusion/exclusion criteria will be specified.

Outcomes:

All variables listed below are relevant and needed to validate the PREMII tool.

The following variables in the NNRD have been identified that are needed to match the factors and the associated levels of the PREMII:

- Gestational age
- Sex
- Body weight (daily)
- At admission:
    - Percent (%) of oxygen saturation
    - Temperature
    - Clinical diagnosis
- Respiratory support
    - Use and type of respiratory interventions (daily)
    - Mode of respiratory support via endotracheal tube (daily)
    - Respiratory rate at admission (per min); use of nitric oxide, chest drain, and surfactant (daily)
    - Receiving a continuous infusion of a pulmonary vasodilator (daily)
- Use of any supplementary medications
- Method of feeding (daily)

The following list of variables is also included in the PREMII tool, but is not available in the database. A modified PREMII score will be calculated based on the subset of data elements that are captured in the NNRD (listed above).

- Percentage of oxygen administration
- Apnea, bradycardia, and desaturation events
- Thermoregulation
- Duration of feeding or infant problems during feeding

In addition to the variables needed to match the PREMII tool, LOS will also be examined in the NNRD.

- Discharge
    - Minutes from birth
    - Reason

Study Size:

The sample size for the current study will depend upon the total number of eligible premature infants <28 weeks gestation born after 2012 who are available for analysis in the NNRD. The total sample size is estimated to be approximately 10,000 premature infants based on a preliminary count of 2,595 premature infants born in 2016 alone.

Data analysis:

Upon receipt of the NNRD data, the analytic team will convert the data into SAS V9 format. Two analytic datasets will be created: the baseline dataset and the longitudinal dataset. The baseline dataset will be structured with one record per patient, and will contain all the single-observation baseline data such as gestational age and birth weight. The longitudinal dataset will contain all repeated measures such as the variables that are associated with the PREMII factors and levels that are measured and collected daily.

The baseline characteristics of the study population will be described at birth with respect to each of the explanatory variables in the NNRD that are components of the PREMII. Descriptive statistics will be presented for the study population overall and stratified by gestational week at birth. Descriptive data will primarily consist of counts and percentages for data that is categorical in nature, and measures of central tendency, mean or median, and dispersion, standard deviation, range or inter-quartile range, for continuous variables.

A modified baseline PREMII score will be calculated for each premature infant using the NNRD measurements taken at week 36 PMA and at week 40 PMA.

The predictive validity of the modified PREMII score measured at week 36 of PMA to predict subsequent LOS will be evaluated using a Cox proportional hazards regression model. The predictive validity of each explanatory variable that is a component of the modified PREMII score also will be evaluated as a predictor of LOS using Cox proportional hazards regression models. Infants who died in the NNU will be censored from the analysis of LOS. An exploratory analysis will also be conducted to describe the distribution of infants at each level of each factor in the PREMII score by week of PMA to explore changes in infant's functional status during their stay in the NICU.

The discriminant validity of the modified PREMII scored measured at week 36 of PMA to predict discharge from the NNU during that same week (i.e., week 36) will be evaluated using logistic regression models. If the week 36 modified PREMII score is missing for an infant, then the week 35 modified PREMII score will be used in the analysis (last observation carried forward). Similarly, for the exploratory analysis, the discriminant validity of the modified PREMII score measured at week 40 of PMA to predict discharge from the NICU during that same week (i.e., week 40 +1 to 6 days) will be evaluated using logistic regression models. If the week 40 modified PREMII score is missing for an infant, then the week 39 modified PREMII score will be used in the analysis (last observation carried forward).

Milestones:

The specific study milestones are dependent upon the time for data request, research ethics submission and approval, and receipt of NNRD data. The following milestones are based on the assumption that the requested NNRD data would be received by August 30, 2018.

The planned deliverables for this study and anticipated date of completion include:

- Final study protocol (June 29, 2018);
- Final statistical analysis plan (June 29, 2018);
- Results tables (October 5, 2018); and
- Final study report (November 2, 2018).

5. AMENDMENTS AND UPDATES

Not applicable**.**

6. MILESTONES

Milestone	Planned Date
Date (range) of data collection:	30 August 2018 (anticipated date of receipt of data from dataset holders)
Study progress report(s):	N/A
Final report:	2 November 2018

7. RATIONALE AND BACKGROUND

Approximately four million babies are born every year in the United States and about 11% of those are born prematurely. Similarly, around 60,000 babies are born prematurely in the United Kingdom every year and one in nine babies born in England will spend time in the neonatal unit (NNU).[1-3] Given that caring for extremely premature infants in the NNU can pose a heavy resource burden to the health care system, being able to identify infants that are physically ready to be discharged from the NNU is critical in reducing resource utilizations and hospital expenditures.

Reducing length of stay (LOS) in the NNU may also be considered a clinically meaningful outcome for extremely premature infants. However, because LOS is not necessarily solely related to an infant's overall health or functional status, and many factors such as a caregiver's readiness to care for the infant at home can also impact LOS. As an alternative to LOS, Shire developed the PREMature Infant Index (PREMII), a clinician reported outcome, to assess overall functional status of extremely premature infants in the NNU at 36 weeks of postmenstrual age (PMA), defined as gestational age plus chronological age (time elapsed from birth)[4]. Functional status is defined as what the infant can do with respect to key functional areas, as a reflection of the infant's overall health and maturation. Specifically, the PREMII evaluates 8 factors: respiratory support, oxygen administration, apnea, bradycardia, desaturation, thermoregulation, feeding, and weight gain, each scored on between 3–6 levels of functioning. It is expected that the PREMII score would be associated with an infant's physical readiness for discharge as well as LOS in the NNU. As the PREMII is intended to be used as a secondary clinical endpoint in an upcoming phase IIb/III clinical trial, it is important to conduct a validation of this tool using real-world data to help support the use of the PREMII as a clinical endpoint.

In 2016, a feasibility study was conducted to identify potential real-world databases that can be used as a validation data source for the PREMII. From this feasibility study, the United Kingdom's National Neonatal Research Database (NNRD) was selected as the most appropriate data source to validate the PREMII. The following sections describe the details of this validation study.

8. RESEARCH QUESTION AND OBJECTIVES

8.1 Research Question

Shire is interested in validating the PREMII tool against a real-world database to examine the association of the PREMII with an infant's functional status in the NNU using real-world practice patterns.

8.2 Objectives

The overarching aim of the study is to validate the PREMII tool in a representative sample of premature infants in a real-world practice setting.

The primary objective of this study is to evaluate the predictive validity of the PREMII score calculated from factor levels measured at 36 weeks of PMA with subsequent LOS in the NICU.

The secondary objective is to assess the discriminate validity of the modified PREMI score measured at 36 weeks of PMA with discharge (yes/no) from the NNU during that same week.

The exploratory objectives are as follows:

1. Describe the distribution of infants at each level of each factor in the PREMII score by week of PMA.
2. Evaluate the discriminate validity of the PREMII score measured at 40 weeks of PMA with discharge from the NICU (yes/no) during that same week.

9. RESEARCH METHODS

9.1 Study Design

A retrospective database analysis will be conducted. Patients will be identified within the NNRD.

The design chosen for this study allows for the analysis of real-world data already routinely collected in the NNU to examine the validity of the PREMII as a measure of an infant's functional status.

A national research ethics committee application will also be submitted and an approval from the ethics committee is needed before data can be released from the data holder.

9.1.1 Data Collection

Infants identified for inclusion from the NNRD will be retrospectively followed from birth to discharge from participating NNUs.

9.2 Setting

9.2.1 Number and Type of Patients

The source population will be all infants admitted to a NNU and recorded in the NNRD. A sample of infants meeting the criteria outlined in Section 9.2.3 and having data on or after 2012 in the NNRD will be drawn from the source population and will be included in the study. Please see Section 9.4 for a detailed description of the NNRD.

9.2.2 Sites and Regions

All NNUs in England provide data in the NNRD will be invited to allow inclusion of their data in the study.

9.2.3 Selection Criteria

9.2.3.1 Inclusion Criteria

Eligible infants need to have complete data on demographics (i.e. month of birth, year of birth, birth weight, gestational age and sex) and admission and discharge details (i.e. date and time derived from minutes from birth) in order to be included in the study.

Infants must have been less than 28 weeks of gestational age upon admission to the NNU after birth.

9.2.3.2 Exclusion Criteria

Infants with congenital heart disease, congenital diaphragmatic hernia, or congenital malformations are not eligible for the study.

Congenital heart disease for this study includes having bulbus cordis anomalies of cardiac septal closure (excluding patent foramen ovale), or other congenital anomalies of heart or of circulatory system (excluding patent ductus arteriosus).

Congenital malformation for this study includes having congenital musculoskeletal deformities of chest, congenital malformations of lung or congenital malformations of upper airway.

The specific ICD codes to indicate these congenital conditions are summarized below:

Table 9.1: ICD-9-CM and ICD-10-CM Codes to Identify Congenital Heart Disease, Diaphragmatic Hernia and Congenital Malformations

Comorbidities		ICD-9-CM Code(s)	ICD-10-CM Code(s)
Congenital heart diseases	Bulbus cordis anomalies and anomalies of cardiac septal closure (excluding patent foramen ovale)	745.1 - 745.4, 745.6 - 745.9	Q20, Q21.0, Q21.2 - Q21.9
	Other congenital anomalies of heart	746	Q22, Q23, Q24
	Other congenital anomalies of circulatory system (excluding patent ductus arteriosus)	747.1 - 747.9	Q25.1 - Q25.9, Q26, Q27, Q28
Diaphragmatic hernia		756.6	Q79.0, Q79.1
Congenital malformations	Other congenital musculoskeletal deformities of chest	754.89	Q67.8
	Congenital malformations of lung	748.4 - 748.6	Q33
	Congenital malformations of upper airway	748.0 - 748.3	Q30, Q31, Q32

9.3 Outcomes

The outcomes selected for this study were selected specifically to provide data relating to the individual factors/levels of the PREMII tool.

Data for the following variables, which have been identified in the NNRD using the field list in the Neonatal Data Set[5], will be extracted for all eligible infants in this study:

- Gestational age
- Sex
- At admission:
    - Percent (%) of oxygen saturation
    - Temperature
    - Clinical diagnosis
- Body weight (daily)
- Respiratory support
    - Use and type of respiratory interventions (daily)
    - Mode of respiratory support via endotracheal tube (daily)

  - Respiratory rate at admission (per min); use of nitric oxide, chest drain, and surfactant (daily)
  - Receiving a continuous infusion of a pulmonary vasodilator (daily)

- Use of any supplementary medications
- Method of feeding (daily)

The following list of variables is also included in the final PREMII tool, but is not available in the database. A modified PREMII score will be calculated based on the subset of data elements that are captured in the NNRD (listed above).

- Percentage of oxygen administration
- Apnea, bradycardia, and desaturation events
- Thermoregulation
- Duration of feeding or infant problems during feeding

In addition to the variables needed to match the PREMII tool, LOS will also be examined in the NNRD.

- Discharge
  - Minutes from birth
  - Reason

9.4 Data Sources

The NNRD will be used to obtain relevant data on infants during their stay in the NNU. The NNRD is a population level database that currently holds data on over 800,000 infants and is updated each quarter. All 200 NNUs in England, Wales and Scotland contribute data to the NNRD. Data from the NNRD originate from information entered by clinicians and nursing staff onto the Bager.net platform at the point of care. Data are extracted from these electronic patient records at intervals, undergo quality assurance procedures, are anonymized and entered into the NNRD by the Neonatal Data Analysis Unit.

The study requires confirmation of agreement to participate from all NNU leads participating in the NNRD. Data from NNU leads choosing not to participate in this study will not be included in the data extract.

9.5 Study Size

The sample size for the current study will depend upon the total number of eligible premature infants who are available for analysis in the NNRD. The total sample size is estimated to be approximately 10,000 premature infants based on a preliminary count from the NNRD of 2,595 premature infants born in 2016 alone.

9.6 Data Analysis

Upon receipt of the NNRD data, the analytic team will convert the data into SAS V9 format. Two analytic datasets will be created: the baseline dataset and the longitudinal dataset. The baseline dataset will be structured with one record per patient, and will contain all the single-observation baseline data such as gestational age and birth weight. The longitudinal dataset will contain all repeated measures such as the variables that are associated with the PREMII factors and levels that are measured and collected daily.

The baseline characteristics of the study population will be described at birth with respect to each of the explanatory variables in the NNRD that are components of the PREMII. Descriptive statistics will be presented for the study population overall and stratified by gestational week at birth. Descriptive data will primarily consist of counts and percentages for data that is categorical in nature, and measures of central tendency, mean or median, and dispersion, standard deviation, range or inter-quartile range, for continuous variables.

A modified PREMII score will be calculated for each premature infant based on the variables in the NNRD which provide data for the respective factors and levels of the PREMII, as stated in Section 9.3. The modified PREMII score will use the NNRD measurements taken at week 36 PMA (the first daily record recorded in the 36^{th} week PMA), and at week 40 PMA.

The predictive validity of the modified PREMII score measured at week 36 of PMA in the NICU to predict subsequent LOS will be evaluated using a Cox proportional hazards regression model. The predictive validity of each explanatory variable that is a component of the modified PREMII score also will be evaluated as a predictor of LOS using Cox proportional hazards regression models. Infants who died in the NNU will be censored from the analysis of LOS. The statistical significance of the modified PREMII score and each of its components in predicting of LOS will be determined using a Wald test criterion for the parameter estimate of $P<0.05$. An exploratory analysis will also be conducted to describe the distribution of infants at each level of each factor in the PREMII score by week of PMA to explore changes in infant's functional status during their stay in the NICU.

The discriminant validity of the modified PREMII score measured at week 36 PMA to predict discharge from the NNU during that same week (i.e., week 36) will be evaluated using logistic regression models. If the week 36 modified PREMII score is missing for an infant, then the week 35 modified PREMII score will be used in the analysis (last observation carried forward). Infants with imputed week 36 modified PREMII scores will be removed in a sensitivity analysis. The statistical significance of higher PREMII score predicting discharge

at week 36 will be determined using a Wald test criterion for the parameter estimate of $P<0.05$. Similarly, for the exploratory analysis, the discriminant validity of the modified PREMII score measured at week 40 of PMA to predict discharge from the NICU during that same week (i.e., week 40 +1 to 6 days) will be evaluated using logistic regression models. If the week 40 modified PREMII score is missing for an infant, then the week 39 modified PREMII score will be used in the analysis (last observation carried forward).

9.7 Data Management

Upon receipt of data from the NNRD, data will be stored according to the specifications of the data stewards, and archived for a period of time determined by them, but not to exceed seven years. It is anticipated that all raw data and associated analytic datasets will be stored by ICON on a secured server, with restricted access to relevant team members only.

9.8 Limitations of the Research Methods

The main challenge with conducting this validation study is that there are some factors/levels in the PREMII that are not captured as variables in the NNRD. Although a feasibility assessment was conducted and identified the NNRD as the real-world database which would be the most appropriate validation source for the PREMII, there are limitations in using a retrospective database where the data routinely collected do not necessarily translate exactly to the variables of a tool. However, to address this limitation a modified PREMII score will be calculated based on the subset of data elements that are captured in the NNRD. The implications of this approach will be discussed and explored at the analysis and reporting stage.

In addition, LOS is confounded by numerous factors not directly related to the infant's overall health or functional status measured using the PREMII. An evaluation of baseline characteristics will be conducted and key covariates such as geographic location of the NNU may be adjusted during the analysis stage to reduce confounding.

9.9 Informed Consent

Given that anonymized patient data from the NNRD will be reviewed for this study, no informed consent will be required. However, NNUs with data in the NNRD will need to agree to participation of having their data being used for this study.

9.10 Privacy and Confidentiality

Only anonymized data will be requested during data request. ICON Clinical Research Limited will be receiving patient-level data files from the participating NNUs. No patient-level data extract files will be shared or disclosed to non-study stuff or to the study sponsor.

10. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

The outcomes of the study will mainly be used internally to support the use of the PREMII as a clinical endpoint. Findings may be presented at conferences or in a published paper.

10.1 Public Posting of Study Information

Not applicable

10.2 Submission of Summary of Final Study Report to Competent Authorities of Member States Concerned and Ethics Committees

Not applicable

10.3 Publications Policy

Unless otherwise required by the journal in which the publication appears, or the forum in which it is made, authorship will comply with the International Committee of Medical Journal Editors (ICMJE) current standards. Participation in this study does not confer any rights to authorship of publications.

11. REFERENCES

1. Temple MW, Lehmann CU, Fabbri D. Predicting Discharge Dates From the NICU Using Progress Note Data. Pediatrics 2015;**136**(2):e395-405.
2. BabyCentre Medical Advisory Board. The neonatal unit. Secondary The neonatal unit 2016. https://www.babycentre.co.uk/a555453/the-neonatal-unit.
3. Premature Babies. Everything you need to know about premature babies. Secondary Everything you need to know about premature babies 2017. https://www.premature-babies.co.uk/.
4. Engle WA. Age terminology during the perinatal period. Pediatrics 2004;**114**(5):1362-4.
5. Imperial Colledge. Neonatal Data Set (ISB1595). Secondary Neonatal Data Set (ISB1595). http://www.imperial.ac.uk/media/imperial-college/medicine/dept-medicine/infectious-diseases/neonatology/Neonatal-dataset-ISB1595-release-1-version-22.pdf.

APPENDICES

APPENDIX 1 Protocol History

Document	Date	Global/Country/Site Specific
Original Protocol	10 Nov 2017	United Kingdom

{***Every protocol will have this appendix. Upon adding amendments, the above table will be hyperlinked to the beginning of the protocol amendment table which provides the changes and rationale. Note: Typographical errors need not be identified.***}

Protocol Amendments		
Summary of Change(s) Since Last Version of Approved Protocol		
Amendment Number **#**	**Amendment Date** **DD Mon YYYY**	**Global/Country/Site Specific** **Global**
Description of Change		**Section(s) Affected by Change**
{***Insert a high-level summary of the change. Include rationale, if applicable. Repeat this row as needed for each change.***}		Section #.# Section #.#.#