Prospective identification and causality evaluation of suspected adverse drug reactions in neonates

Running title: Evaluating ADRs in neonates

Eve K. Roberts[1], Dan Hawcutt[1,2]*, Mark A. Turner[1,3]*

The authors confirm that the PI for this paper is Mark A. Turner and that he had direct clinical responsibility for patients

1: Department of Women's and Children's Health, Institute of Translational Medicine, University of Liverpool, UK, a member of Liverpool Health Partners

2: NIHR Alder Hey Clinical Research Facility, Alder Hey Children's Hospital, Liverpool, UK, a member of Liverpool Health Partners

3; Liverpool Women's NHS Foundation Trust, Liverpool, UK, a member of Liverpool Health Partners

*These authors contributed equally to this manuscript

There are no conflicts of interest to declare for any authors.

The authors are happy to collaborate with other research teams using anonymised data, on request. Please contact the corresponding author for further details.

Corresponding author:

Dr Eve Roberts
Institute of Translational Medicine
Department of Child Health
University of Liverpool
First Floor
Liverpool Women's Hospital
Crown Street
Liverpool
L8 7SS
Eve.roberts@ncl.ac.uk
everoberts@doctors.org.uk

Abstract

Neonates experience adverse drug reactions (ADRs), but under-reporting of suspected ADRs to national spontaneous reporting schemes in this population is particularly high. A prospective observational study collected suspected neonatal ADRs at a tertiary neonatal unit. Cases were analysed for causality by six assessors using three existing methods. Sixty-three suspected ADR cases were identified in 35/193 neonates (18.1%). The proportion of suspected ADRs where the drug was prescribed 'off-label' was 30/68 (44.1%). When 34 cases were assessed for causality using three methods, global kappa scores of less than 0.3 for each tool suggested only 'fair' inter-rater reliability. Neonatal ADRs can be captured and occur from a variety of drugs affecting many organ systems. The current tools for assessing causality need to be adapted before they can reliably assess neonatal ADRs.

What is already known about this subject:

- Neonates experience adverse drug reactions (ADRs), but there has been little prospective evaluation of the drugs suspected, or the causality
- Relatively few spontaneous reports of suspected ADRs relate to neonates, hindering pharmacovigilance in this population
- Several methods are currently available for assessing causality of suspected ADRs in neonates, but comparative data are limited

What this study adds:

- Suspected ADRs were observed to affect 18% of neonatal inpatients, affecting most neonatal organ systems
- A wide range of drugs were observed to cause suspected ADRs, with gentamicin, morphine, and dopamine being most frequently implicated
- Current adverse drug reaction causality assessment methods exhibit low inter-rater reliability, and further development of these methodologies in this population is required

Introduction

Approximately 15 million babies are born premature each year globally[1]. Many of these babies will be prescribed medicines, and yet most previous pharmacovigilance studies have omitted part or all the inpatient neonatal population from their work. An adverse drug reaction is defined by the World Health Organisation as 'a response to a drug which is noxious and unintended, and which occurs at doses normally used in man' [2]. Recent studies into adverse drug reactions in children have indicated a considerable health risk for this population, with incidence rates ranging from 0.4% to 10.3% for paediatric hospital admissions related to ADRs and 0.6% to 16.8% for the proportion of children experiencing an ADR during their admission [3].

Neonates are subject to different adverse drug reaction profiles in comparison to older children and adults [4,5]. The development of a child from conception to adulthood is dynamic, and changes in organ function and body composition affect pharmacodynamics and pharmacokinetics [6]. Neonates born preterm are subject to further variation in drug absorption, distribution, metabolism and excretion. A study conducted in a neonatal intensive care unit reported that 29.6% of neonates received more than four medications and 7.6% received 10 or more [7]. Further studies show that up to 90% of inpatient neonates receive off-label or unlicensed medications, a considerable risk factor for developing an ADR [7,8]. Some medications are uniquely harmful to neonates [5]. ADR reporting rates to spontaneous reporting schemes for children are low and neonates are particularly poorly represented [9,10]. However, the historical lack of inclusion of neonates in drug trials means that it is particularly important to generate pharmacovigilance data in this population [5].

Evaluation of ADR reports is also very important. Evaluating the severity of an ADR assesses the importance of an ADR in a clinical context, and a neonatal adverse event severity scale has recently been developed through a Delphi consensus approach[11]. Causality assessment tools enable structured assessment of the likelihood of drug-reaction accountability and could help to reduce disagreements between clinicians, thus increasing ADR reporting. Such tools are used by regulatory agencies for the evaluation of ADR reports [12].

The Naranjo algorithm is a widely used causality assessment method [13]. However, a large-scale observational study into ADRs in children concluded that the Naranjo algorithm was not suitable to assess paediatric ADRs [12]. Consequently, a new tool was developed for children and showed greater inter-rater reliability [12]. The resulting 'Liverpool ADR Causality Assessment Tool' (LCAT) has not been assessed in neonates. A neonatal modification of the Naranjo algorithm was developed recently in one centre, but this has not been validated in another site [14]. This study aims to

compare methods for evaluating adverse drug reactions in neonates, in order to determine which method, if any, is reliable to use for ADR evaluation in this population.

Methods

Study design and participants

A prospective observational study, the Adverse Drug Reactions in Neonates (ADRIN) study, was undertaken at a tertiary neonatal centre in the UK. All neonatal inpatients were monitored daily for nine weeks. Neonates were reviewed daily up to 28 days post-term (corrected gestational age). Suspected neonatal ADR cases were identified by medical or nursing teams or by the researcher, a 5th year medical student. The review process is outlined in the daily structured clinical review guidance found in appendix 2. All suspected ADRs were discussed and approved for inclusion by the principal investigator, a consultant neonatologist.

ADR case causality assessment

Figure 1 outlines case selection for causality assessment. ADRs caused by drugs used by parents (i.e. by mother in labour) were excluded as the causality assessment tools used have not been validated for assessing this subtype of ADR. Six assessors (Table S1) were asked to complete three known causality assessment assessments for each case; the Karch and Lasagna method, the New Adverse Drug Reactions algorithm for Infants in the Neonatal Intensive Care Unit (referred to within as the Du Lehr method) and the Liverpool ADR Causality Assessment Tool [12, 14, 15].

Statistical analysis of ADR reports

Neonatal ADR data was summarised (Tables S2-6). All suspected ADR cases were included in the ADR case analysis.

Statistical analysis of causality assessments

Inter-rater reliability:

Inter-rater reliability was calculated using non-weighted, weighted and global kappa scores. Percentage exact agreement and percentage extreme disagreement were also calculated to show the level of concordance between pairs of assessors.

Inter-tool reliability:

Inter-tool reliability was calculated using kappa scores to measure agreement between the ratings for the same cases assessed by the same assessor using two different tools.

The level of kappa acceptability for both inter-rater and inter-tool reliability was chosen to match that used in the 'Adverse Drug Reactions in Children' (ADRIC) research programme: <0.2=poor, 0.21-0.40=fair, 0.41-0.60=moderate, 0.61-0.80=good, 0.81-1.00=very good [12,16].

Results

Over the data collection period, a total of 193 neonates were inpatients on the neonatal unit. Sixty-three reports detailing suspected ADRs were recorded during the data collection period. Fifty-six reports were ascribed to drugs prescribed for the neonate, and seven from maternal drugs.

Neonatal characteristics

The gestational ages of the neonates at birth ranged from 23 + 6 weeks to 40 + 4 weeks. The neonates' corrected gestational ages at time of experiencing an ADR ranged from 26 + 1 weeks to 40 + 5 weeks (median 33 + 4 weeks). The birth weights of the neonates ranged from 570g to 3990g, with a mean birth weight of 1874g. The neonates' working weights at time of ADR reporting ranged from 580g to 3990g (mean 1390g).

Adverse Drug Reactions Identified

Thirty five of the 193 neonates reviewed (18.1%) were suspected to have experienced one or more ADRs. Of these neonates, 28 (80%) were affected by drugs prescribed to the neonate, and seven (20%) by maternal drugs. The number of ADR reports per neonate ranged from 1-6 (median 1 report). Table 1 summarises the medications most commonly reported, and reactions suspected.

Table S2 summarises the suspected ADR reports captured. The number of suspected ADRs exceeds the number of reports as some reports detailed more than one reaction. Thirty-six reports contained only one suspected medication, 18 reports listed two medications, while two reports contained three suspected medications. Overall 31 different drugs were suspected to have caused neonatal ADRs.

The most commonly reported suspected ADRs were pyrexia (n=4), tachycardia (4), thrombocytopenia (3), altered consciousness (3) and renal failure (3). A complete list of the suspected ADRs identified is shown in table S3. The most common drug groups (by ATC classification) causing ADRs (not including those ADRs suspected to be from maternal drugs) were those drugs in the cardiovascular system group (28), the anti-infectives for systemic use group (22) and the nervous system group (9) (table S4). Thirty ADR reports contained medications that had been prescribed to the neonate off-label (table S5).

Seven reports were from medications administered to the mother, either during pregnancy or labour, identifying seven suspected ADRs. The reports detailed a total of 11 suspected drugs, nine prescribed to the mother in pregnancy, two prescribed to the mother in labour and one report of illicit drug use (table S6).

Causality assessments

Six assessors undertook each of the three assessments on 34 different cases, resulting in 612 total assessments (Table S7). A chi squared test showed that the excess of definite ratings using the Du Lehr method was highly statistically significant, p-value < 0.001.

Inter-rater reliability

Pair-wise kappa scores were measured between all six assessors using each of the three tools (tables S8-10). Weighted kappa scores ranged from 0.148 to 0.454 for the Karch and Lasagna algorithm, 0.114 to 0.483 for the Du Lehr and 0.121 to 0.428 for the LCAT. Most weighted kappa scores for each pair-wise comparison for each tool corresponded to 'fair' inter-rater reliability (Supplementary tables S8-S10). Percentage exact agreement between the ratings given to each case by each of two assessors, ranged from 14.7% to 41.2% for the Karch and Lasagna algorithm, 38.2% to 58.8% for the Du Lehr tool and 26.5% to 61.8% for the LCAT. Extreme disagreement ranged from 8.82% to 35.3% for the Karch and Lasagna algorithm, 11.8% to 38.2% for the Du Lehr tool and 0% to 17.6% for the LCAT (Supplementary tables S8-S10). Global kappa scores were measured to outline the inter-rater agreement between all six assessors. These were 0.157 (CI 0.0741 – 0.239) for the Karch and Lasagna algorithm, 0.254 (CI 0.139 – 0.369) for the Du Lehr tool and 0.209 (CI 0.121 – 0.297) for the LCAT.

Inter-tool reliability

Kappa scores were also calculated to measure inter-tool reliability, the agreement when each assessor used different tools to assess the same 34 cases (Table 2). The highest kappa score was seen when comparing each assessor's ratings using the Karch and Lasagna algorithm and the LCAT.

Discussion

In this study ADRs were suspected to have affected term and preterm neonates, involving many different drugs. It can be challenging to distinguish between ADRs and Adverse Events (AEs), and this study demonstrates opinion contributes to categorisation of causality. An effective causality assessment tool would help to translate clinician concern into categorical likelihood, aiding clinical drug therapy risk-benefit analysis and ADR reporting, ultimately improving neonatal care.

This study identified 63 suspected neonatal ADRs over a nine-week period, which would produce approximately 350 per year in a single unit. Nationally in the UK, only 97 were reported to the Yellow Card spontaneous reporting scheme in a 10-year period [10]. The definition of an ADR used in this study allowed for the inclusion of ADRs to drugs prescribed unlicensed or off-label, a practice which commonly occurs in neonatology [17].

A wide range of reaction types were observed and there was at least one ADR report for nearly every organ system. While cardiovascular system drugs, anti-infectives for systemic use, nervous system drugs, and sensory organ drugs were the most commonly identified ATC categories, the limited size of this study means it is not possible to determine whether this is because they are the more harmful, or more commonly prescribed, or both. A recent quasi-systematic review outlined the most commonly prescribed drugs in NICUs worldwide [18]. It found nine of the top twenty most cited drugs were also listed on the A-PINCH list, a list of medications that pose high risks if medication errors occur [19]. The list includes anti-infectives, potassium and concentrated electrolytes, insulin and narcotics and sedatives, all of which are used on neonatal units. In total 22 of the 78 reported drugs in this study are A-PINCH listed drugs [18,19].

In recent years, it has become apparent that many pre-existing ADR assessment tools are inappropriate for assessing paediatric ADRs. The results of evaluating the three causality assessment methods to determine their appropriateness for assessing neonatal ADRs show no clear optimum method. The highest Kappa scores demonstrated inter-rater reliability of less than 50%, suggesting even the best performing tool could not yet have a useful clinical implementation.

The highest inter-tool reliability was observed between the Karch and Lasagna algorithm and the LCAT (table 2). However, this only shows that these tools most often lead the user to the same outcome, regardless of whether this outcome is an over or under-estimate of causality. This suggests that it is important to be consistent in causality assessment methodology within pharmacovigilance studies, whilst consistent methods of causality assessment across pharmacovigilance studies would facilitate comparison between results. Previously published studies of ADRs have only reported on

those deemed probable or definite using varying causality assessment methods, demonstrating that the method of causality assessment used affects the incidence rates and results reported, and thus the interpretations made.

Some assessors demonstrated moderate and good intra-rater reliability, however, their inter-rater reliability scores were below average. This suggests that the assessors with good intra-rater reliability were consistent in their reasoning, but they may have been evaluating different concepts to other reviewers. Further definition of terms within tools and guidance on their use may help to avoid this.

This study demonstrates that neonatal ADRs can be captured, but that more work is needed to design reliable causality assessment tools. The improvement in inter-rater reliability seen when using a neonate-specific method, suggests that a population-specific method could be advantageous. A full evaluation of such ADRs will also require severity and avoidability assessments, and future research focussing on these areas will help to bring pharmacovigilance in neonates in line with that of older populations.

Acknowledgements

The authors would like to thank all the neonatal unit staff who contributed to the results of the study, Dr Jamie Kirkham for his help with the statistical analysis, and all the patients and parents who were resident on the unit throughout the study period.

This is a summary of independent research funded by University of Liverpool and carried out at the National Institute for Health Research (NIHR) Alder Hey Clinical Research Facility. The views expressed are those of the author(s) and not necessarily those of the University of Liverpool, NHS, the NIHR or the Department of Health.

References

1. Bliss Charity. Bliss website: The global picture https://www.bliss.org.uk/research-campaigns/research/neonatal-care-statistics/the-global-picture. Accessed January 13, 2020.

2. World Health Organisation website. Safety of Medicines- A Guide to Detecting and Reporting Adverse Drug Reactions- Why Health Professionals Need to Take Action.; 2002. http://apps.who.int/medicinedocs/en/d/Jh2992e/13.html. Accessed September 14, 2016.

3. Smyth RMD, Gargon E, Kirkham J, et al. Adverse drug reactions in children-A systematic review. PLoS One. 2012;7(3):e24061. doi:10.1371/journal.pone.0024061.

4. Yaffe SJ, Estabrrok RW. Rational Therapeutics for Infants and Children. [Electronic Book] : Workshop Summary University of Liverpool Library. Washington, D.C: National Academy Press; 2000.

5. Hawcutt DB, O'Connor O, Turner M a. Adverse drug reactions in neonates: could we be documenting more? Expert Rev Clin Pharmacol. 2014;7(6):807-820. doi:10.1586/17512433.2014.956090.

6. Mahmood, I. (2016). Developmental pharmacology: impact on pharmacokinetics and pharmacodynamics of drugs. In *Fundamentals of Pediatric Drug Dosing* (pp. 23-44). Adis, Cham.

7. Aranda J V., Cohen S, Neims AH. Drug Utilization in a Newborn Intensive Care Unit. J Pediatr. Vol 89. Mosby; 1976. doi:10.1016/S0022-3476(76)80478-7.

8. Turner S, Nunn AJ, Fielding K, Choonara I. Adverse drug reactions to unlicensed and off-label drugs on paediatric wards: A prospective study. Acta Paediatr Int J Paediatr. 1999;88(9). doi:10.1080/08035259950168469.

9. Hawcutt DB, Mainie P, Riordan A, Smyth RL, Pirmohamed M. Reported paediatric adverse drug reactions in the UK 2000-2009. Br J Clin Pharmacol. 2012;73(3). doi:10.1111/j.1365-2125.2011.04113.x.

10. Hawcutt DB, Russell N-J, Maqsood H, et al. Spontaneous Adverse Drug Reaction Reports for Neonates and Infants in the UK 2001-2010: Content and Utility analysis. Br J Clin Pharmacol. 2016. doi:10.1111/bcp.13067.

11. Salaets, T., Turner, M. A., Short, M., Ward, R. M., Hokuto, I., Ariagno, R. L., ... & Roberts, E. (2019). Development of a neonatal adverse event severity scale through a Delphi consensus approach. *Archives of Disease in Childhood*, *104*(12), 1167-1173.

12. Gallagher RM, Kirkham JJ, Mason JR, et al. Development and inter-rater reliability of the Liverpool adverse drug reaction causality assessment tool. PLoS One. 2011;6(12). doi:10.1371/journal.pone.0028096.[13] Naranjo CA, Busto U, Sellers EM (1981) A method for estimating the probability of adverse drug reactions. Clinical Pharmacology and Therapeutics 30: 239–245

14. Du W, Lehr VT, Lieh-Lai M, et al. An algorithm to detect adverse drug reactions in the neonatal intensive care unit. J Clin Pharmacol. 2013;53(1):87-95. doi:10.1177/0091270011433327.

15. Karch FE, Lasagna L. Toward the operational identification of adverse drug reactions. Clin Pharmacol Ther. 1977;21(3):247-254. Doi:0.1002/cpt1977213247

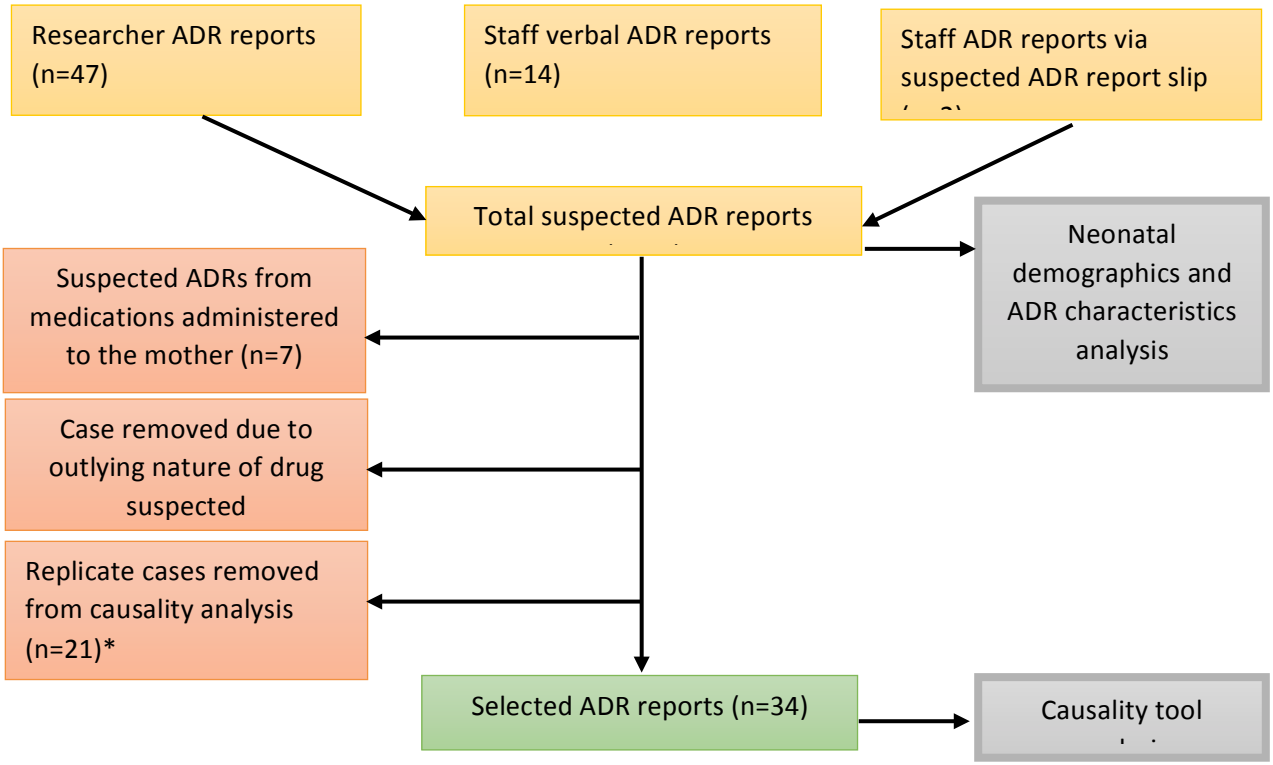
16. Altman DG. Practical Statistics for Medical Research. Chapman and Hall; 1991.

17. Belén Rivas A, Arruza L, Pacheco E, Portoles A, Diz J, Vargas E. Adverse drug reactions in neonates: a prospective study. Arch Dis Child. 2016;101(4):371-376. doi:10.1136/archdischild-2015-309396.

18. Krzyżaniak N, Pawłowska I, Bajorek B. Review of drug utilization patterns in NICUs worldwide. J Clin Pharm Ther. 2016;41(6):612-620. doi:10.1111/jcpt.12440.

19. South Australia Health Website: Medicines and Drugs. High risk medicines. https://www.sahealth.sa.gov.au/wps/wcm/connect/Public+Content/SA+Health+Internet/Clinical+resources/Clinical+topics/Medicines+and+drugs/High+risk+medicines/ Accessed July 2, 2017

Figure S1 Suspected ADR case selection for analyses



*Case type replicates	Replicates removed (n)
Gentamicin and renal impairment	3
Antibiotics and watery stoma losses/stool	2
Morphine and altered consciousness	2
Prostaglandin and pyrexia	2
Midazolam/vecuronium and altered urine output	2
Diuretics and hyponatraemia	1
Furosemide and electrolyte disturbance	1
Sodium supplements and hypernatraemia	1
Sodium feredetate and constipation	1
Inotropes and tachycardia	1
Morphine and respiratory depression	1
Gentamicin and thrombocytopenia	1
Benzylpenicillin and thrombocytopenia/leucopenia	1
Inotropes and hypertension/increased urine output	1
Hydrocortisone/inotropes and hypertension/pulmonary haemorrhage	1

Table 1 – Drugs suspected of causing ADRs and the reaction types identified

Drug Suspected of causing ADR (number of reports)	**ADRs identified (number of reactions)**
Gentamicin (8)	Renal Impairment (4)
	Loose stoma output (2)
	Thrombocytopenia (2)
Morphine (6)[a]	Altered consciousness (4)
	Respiratory depression (1)
	Altered consciousness and respiratory depression (1)
Dopamine (5)[a]	Tachycardia (2)
	Hypertension (1)
	Hypertension and increased urine output (1)
	Hypertension and pulmonary haemorrhage (1)
Benzylpenicillin (4)	Thrombocytopenia (1)
	Thrombocytopenia and leucopenia (2)
	Renal impairment (1)
Cyclopentolate eye drops (4)[a]	Tachycardia (1)
	Apnoea (1)
	Vomit (1)
	Desaturation and bradycardia (1)
Dobutamine (4)	Tachycardia (2)
	Cerebral haemorrhage (1)
	Hypertension and large urine output (1)
Prostin (4)	Pyrexia (3)
	Apnoea, desaturations and bradycardia (1)
Aciclovir (3)	Extravasation reaction (1)
	Pyrexia (1)
	Diarrhoea (1)
Furosemide (3)[a]	Electrolyte disturbance (2)
	Raised creatinine and electrolyte disturbance (1)
Hydrocortisone (4)[a]	Bloody gastrointestinal aspirates (1)
	Hyperglycaemia (1)
	Hypertension (1)
	Tachycardia (1)
Hydrochlorothiazide (3)[a]	Hyponatraemia (1)
	Electrolyte disturbance (1)
	Neutropenia (1)
Midazolam (3)	Urinary retention (3)
Spironolactone (3)[a]	Hyponatraemia (1)
	Electrolyte disturbance (1)
	Neutropenia (1)

[a] Drugs prescribed to the neonate off-label

Table 2 Inter-tool reliability: red shading- 'poor' inter-tool reliability, orange shading- 'fair' inter-tool reliability, green shading- 'moderate' inter-tool reliability, dark green shading- 'good' inter-tool reliability

Assessor		Karch and Lasagna and Du Lehr	Karch and Lasagna and LCAT	Du Lehr and LCAT
Assessor 1	Kappa score (95% CI)	0.218 (0.087 to 0.350)	0.580 (0.386 to 0.774)	0.319 (0.183 to 0.454)
	Weighted Kappa score	0.460	0.712	0.616
Assessor 2	Kappa score (95% CI)	0.257 (0.109 to 0.405)	0.261 (0.037 to 0.485)	0.255 (0.094 to 0.415)
	Weighted Kappa score	0.383	0.434	0.538
Assessor 3	Kappa score (95% CI)	-0.022 (-0.157 to 0.112)	0.102 (-0.118 to 0.321)	0.058 (-0.108 to 0.223)
	Weighted Kappa score	0.145	0.236	0.298
Assessor 4	Kappa score (95% CI)	0.109 (-0.037 to 0.255)	0.183 (-0.043 to 0.409)	0.030 (-0.071 to 0.130)
	Weighted Kappa score	0.246	0.360	0.239
Assessor 5	Kappa score (95% CI)	0.067 (-0.072 to 0.206)	0.409 (0.184 to 0.635)	0.071 (-0.100 to 0.242)
	Weighted Kappa score	0.227	0.533	0.285
Assessor 6	Kappa score (95% CI)	0.194 (0.021 to 0.367)	0.455 (0.234 to 0.676)	0.165 (0.004 to 0.325)
	Weighted Kappa score	0.348	0.591	0.338