vision, without FHR monitoring.[5] They did not use FHR monitoring because there was no preparation for emergency cesarean delivery with the mother's consent. The American College of Obstetrics and Gynecology Committee opinion on "Non-Obstetric surgery in Pregnancy" stated that "although there are no data to support specific recommendations regarding non-obstetric surgery and anesthesia in pregnancy, it is important for non-obstetric physicians to obtain obstetric consultation before performing non-obstetric surgery, and the decision to use fetal monitoring should be individualized and each case warrants a team approach for optimal safety of the woman and her baby."[28] FHR monitoring facilitates the best possible care for the fetus, especially when the mother wishes to continue the pregnancy or deliver the fetus if fetal asphyxia is suspected. In this line, we were able to operate safely, stably, and successfully for both mother and fetus under monitoring of FHR with the cooperation of obstetrician and anesthesiologist.

Conclusion

FHR monitoring is useful for AVM surgery during pregnancy. CTG is an appropriate method in the third trimester, whereas ultrasonography, using a transesophageal echo probe, can be used in the second trimester. These methods could have a wider application for cerebrovascular surgery during pregnancy.

Conflicts of Interest Disclosure

The authors declare no conflict of interest concerning the materials or methods used in this study or the findings specified in this article.

References

1) Dias MS, Sekhar LN: Intracranial hemorrhage from aneurysms and arteriovenous malformations during pregnancy and the puerperium. *Neurosurgery* 27: 855–865; discussion 865–866, 1990
2) Graf CJ, Perret GE, Torner JC: Bleeding from cerebral arteriovenous malformations as part of their natural history. *J Neurosurg* 58: 331–337, 1983
3) Robinson JL, Hall CS, Sedzimir CB: Arteriovenous malformations, aneurysms, and pregnancy. *J Neurosurg* 41: 63–70, 1974
4) Tuncali B, Aksun M, Katircioglu K, Akkol I, Savaci S: Intraoperative fetal heart rate monitoring during emergency neurosurgery in a parturient. *J Anesth* 20: 40–43, 2006
5) Balki M, Manninen PH: Craniotomy for suprasellar meningioma in a 28-week pregnant woman without fetal heart rate monitoring. *Can J Anaesth* 51: 573–576, 2004
6) Berman MF, Sciacca RR, Pile-Spellman J, Stapf C, Connolly ES, Jr, Mohr JP, Young WL: The epidemiology of brain arteriovenous malformations. *Neurosurgery* 47: 389–396; discussion 397, 2000
7) Fleetwood IG, Steinberg GK: Arteriovenous malformations. *Lancet* 359: 863–873, 2002
8) Horton JC, Chambers WA, Lyons SL, Adams RD, Kjellberg RN: Pregnancy and the risk of hemorrhage from cerebral arteriovenous malformations. *Neurosurgery* 27: 867–871; discussion 871–872, 1990
9) Gross BA, Du R: Rate of re-bleeding of arteriovenous malformations in the first year after rupture. *J Clin Neurosci* 19: 1087–1088, 2012
10) Gross BA, Du R: Hemorrhage from arteriovenous malformations during pregnancy. *Neurosurgery* 71: 349–355; discussion 355–356, 2012
11) Trivedi RA, Kirkpatrick PJ: Arteriovenous malformations of the cerebral circulation that rupture in pregnancy. *J Obstet Gynaecol* 23: 484–489, 2003
12) Salvati A, Ferrari C, Chiumarulo L, Medicamento N, Dicuonzo F, De Blasi R: Endovascular treatment of brain arteriovenous malformations ruptured during pregnancy—a report of two cases. *J Neurol Sci* 308: 158–161, 2011
13) Fukuda K, Hamano E, Nakajima N, Katsuragi S, Ikeda T, Takahashi JC, Miyamoto S, Iihara K: Pregnancy and delivery management in patients with cerebral arteriovenous malformation: a single-center experience. *Neurol Med Chir (Tokyo)* 53: 565–570, 2013
14) Spetzler RF, Martin NA: A proposed grading system for arteriovenous malformations. *J Neurosurg* 65: 476–483, 1986
15) Gross BA, Du R: Surgical and radiosurgical results of the treatment of cerebral arteriovenous malformations. *J Clin Neurosci* 19: 1001–1004, 2012
16) Parer JT, Ikeda T: A framework for standardized management of intrapartum fetal heart rate patterns. *Am J Obstet Gynecol* 197: 26.e1–6, 2007
17) Parer JT, King T, Flanders S, Fox M, Kilpatrick SJ: Fetal acidemia and electronic fetal heart rate patterns: is there evidence of an association? *J Matern Fetal Neona* 19: 289–294, 2006
18) Zuspan FP, Quilligan EJ, Iams JD, van Geijn HP: Predictors of intrapartum fetal distress: the role of electronic fetal monitoring. Report of the National Institute of Child Health and Human Development Consensus Development Task Force. *Am J Obstet Gynecol* 135: 287–291, 1979
19) Rosen MA: Management of anesthesia for the pregnant surgical patient. *Anesthesiology* 91: 1159–1163, 1999
20) Goodman S: Anesthesia for nonobstetric surgery in the pregnant patient. *Semin Perinatol* 26: 136–145, 2002
21) Gianopoulos JG: Establishing the criteria for anesthesia and other precautions for surgery during pregnancy. *Surg Clin North Am* 75: 33–45, 1995

22) Macarthur A: Craniotomy for suprasellar meningioma during pregnancy: role of fetal monitoring. *Can J Anaesth* 51: 535–538, 2004
23) Reitman E, Flood P: Anaesthetic considerations for non-obstetric surgery during pregnancy. *Br J Anaesth* 107(Suppl 1): i72–78, 2011
24) Minielly R, Yuzpe AA, Drake CG: Subarachnoid hemorrhage secondary to ruptured cerebral aneurysm in pregnancy. *Obstet Gynecol* 53: 64–70, 1979
25) Newman B, Lam AM: Induced hypotension for clipping of a cerebral aneurysm during pregnancy: a case report and brief review. *Anesth Analg* 65: 675–678, 1986
26) Ito T, Toyoshima N, Takahashi H, Ozasa A, Nakajima K, Tatsumura M, Teragawa N, Nobuhara H, Iwanaga M: [Study by fetal movement meter of effects of general anesthesia on fetal movement and fetal cardiac rate]. *Obstetrical and Gynecological Practice* 39: 1115–1118, 1990 (Japanese)
27) Immer-Bansi A, Immer FF, Henle S, Spörri S, Petersen-Felix S: Unnecessary emergency caesarean section due to silent CTG during anaesthesia? *Br J Anaesth* 87: 791–793, 2001
28) ACOG Committee on Obstetric Practice: ACOG Committee Opinion Number 284, August 2003: Nonobstetric surgery in pregnancy. *Obstet Gynecol* 102: 431, 2003
29) Horrigan TJ, Villarreal R, Weinstein L: Are obstetrical personnel required for intraoperative fetal monitoring during nonobstetric surgery? *J Perinatol* 19: 124–126, 1999

Address reprint requests to: Kenji Fukuda, MD, Department of Neurosurgery, Faculty of Medicine, Fukuoka University, 7-45-1 Nanakuma, Jonan-ku, Fukuoka, Fukuoka 814-0180, Japan.
e-mail: kefukuda@fukuoka-u.ac.jp

Circulation Journal
Official Journal of the Japanese Circulation Society
http://www.j-circ.or.jp

ORIGINAL ARTICLE
Myocardial Disease

Cardiovascular Events in Pregnancy With Hypertrophic Cardiomyopathy

Hiroaki Tanaka, MD; Chizuko Kamiya, MD, PhD; Shinji Katsuragi, MD, PhD; Kayo Tanaka, MD; Takekazu Miyoshi, MD; Mitsuhiro Tsuritani, MD, PhD; Masashi Yoshida, MD, PhD; Naoko Iwanaga, MD; Reiko Neki, MD, PhD; Jun Yoshimatsu, MD, PhD; Tomoaki Ikeda, MD, PhD

Background: The influence of the physiological circulatory changes during pregnancy on hypertrophic cardiomyopathy (HCM) is unclear. There have been no comprehensive studies of pregnant women with HCM in the Japanese population.

Methods and Results: A total of 27 pregnancies (23 women with HCM) were retrospectively reviewed. A total of 18 cardiovascular events occurred in 13 of the 27 pregnancies (48%), and 13 of these events (76%) were related to arrhythmia. The cardiovascular events tended to occur in the early stage of pregnancy (≈30 gestational weeks) or postpartum. The events related to arrhythmia mainly occurred in the early stage of pregnancy or at approximately 30 gestational weeks. Four pregnancies were terminated because of cardiovascular events. Cardiovascular events occurred in 8 of 9 pregnancies in women on medication before pregnancy (88%), 7 of 10 pregnancies with high CARPREG score (70%), and in 9 of 12 pregnancies with high ZAHARA score (75%).

Conclusions: Cardiovascular events occurred in more than half of the pregnant women complicated with HCM, and the arrhythmia is the most common cardiovascular event. Medication in the pre-pregnancy period, and CARPREG or ZAHARA score ≥1 were identified as risk factors of cardiac events during pregnancy or postpartum. (*Circ J* 2014; **78:** 2501–2506)

Key Words: Arrhythmia; CARPREG score; Hypertrophic cardiomyopathy; Pregnancy; ZAHARA score

Hypertrophic cardiomyopathy (HCM) is a disease that presents as cardiac muscle dilation with asymmetric diversity. The complications of HCM include arrhythmia, left ventricular outflow obstruction, and diastolic and partial systolic dysfunction because of the myocardial thickening. HCM may result in heart failure, thrombosis, atrial and ventricular arrhythmias, and sudden death, but is often asymptomatic. HCM is thought of as a rare disease, but a recent investigation showed a prevalence of approximately 1.8% in Japan, corresponding to an estimated 21,900 patients with HCM in Japan.[1,2] Therefore, HCM may be more common than previously thought, and this is a matter of concern in the context of pregnancy.

Editorial p 2386

There are few reports on pregnancy in women with exacerbated cardiomyopathy, and the perinatal prognosis of this condition is unclear. The available reports include 7 studies of pregnancy with HCM.[3–9] In the first of these studies, which examined 13 pregnancies with HCM, Turner et al found that vaginal birth was not possible in 2 cases because of worsening angina and in 1 because of breathing difficulties.[3] Autore et al identified 98 survivors and 2 deaths during pregnancy among 100 women with HCM (199 pregnancies),[5] giving a maternal mortality of 10 in 1,000 live births (95% confidence interval 1.1–36.2/1,000), which is higher than that in normal pregnancy. An investigation of the morbidity rate in 40 pregnancies with HCM showed deterioration in New York Heart Association cardiac performance (NYHA class) in 1 of 28 women who were asymptomatic before pregnancy, and in 5 of 12 women who were symptomatic, thus indicating that the perinatal prognosis is excellent in patients who are asymptomatic before becoming pregnant.[5] In a comparison of nonpregnant and pregnant (n=23) women with HCM, the incidence of arrhythmia was higher in those who were pregnant (33.3% vs. 13.4%), but

Received May 12, 2014; revised manuscript received June 23, 2014; accepted July 9, 2014; released online August 6, 2014 Time for primary review: 17 days
Department of Perinatology, National Cerebral and Cardiovascular Center, Suita (H.T., C.K., S.K., K.T., T.M., M.T., M.Y., N.I., R.N., J.Y.); Department of Obstetrics and Gynecology, Faculty of Medicine, Mie University, Tsu (H.T., T.I.), Japan
Mailing address: Hiroaki Tanaka, MD, Department of Perinatology, National Cerebral and Cardiovascular Center, 5-7-1 Fujishirodai, Suita 565-8565, Japan. E-mail: h_tanaka@med.miyazaki-u.ac.jp
ISSN-1346-9843 doi:10.1253/circj.CJ-14-0541

Table 1. Background of the 27 Pregnancies in 23 Women With HCM

Case no.	Age (years)	Parity	Complication	HOCM	D-HCM	Medication pre-pregnancy	NYHA class (pre-pregnancy)
1	25	0	TOF	–	–	–	1
2	25	0	No	+	–	Metoprolol, Verapamil	1
3	31	0	ITP	–	–	–	1
4	33	1	ITP	–	–	–	1
5	33	1	–	–	–	–	1
6	32	2	–	–	–	–	1
7	21	0	Bartter syndrome	–	–	–	1
8	32	0	–	+	–	Verapamil	1
9	39	0	–	+	–	–	1
10	30	0	–	+	–	Diltiazem	2
11	33	1	–	–	–	–	1
12	25	0	–	+	–	Mexiletine, Metoprolol	1
13	30	1	–	+	–	Mexiletine, Metoprolol	1
14	33	0	–	–	–	–	1
15	32	0	–	–	–	–	1
16	34	1	–	–	–	–	1
17	32	0	–	–	–	Propranolol	1
18	32	0	–	+	–	–	1
19	33	0	–	–	–	Diltiazem, Enalapril	1
20	34	1	–	–	–	–	1
21	29	0	–	–	–	–	1
22	33	1	–	–	–	–	1
23	31	1	–	–	–	–	1
24	27	0	–	–	–	–	1
25	33	0	–	–	–	Propranolol	1
26	35	1	–	–	–	–	1
27	28	0	–	–	–	Propranolol, Verapamil	1

D-DCM, dilated phase of hypertrophic cardiomyopathy (HCM); HOCM, hypertrophic obstructive cardiomyopathy; LADs, atrial diameter in endsystole; LVEF, left ventricular ejection fraction; LVOTO, left ventricular outflow tract obstruction; MR, mitral regurgitation; NYHA, New York Heart Association.

(Table 1 continued the next page.)

heart failure and cardiac infarction rates did not differ significantly.[6] There were no deaths in either group, and pregnancy was assumed to have had no influence on the natural course of HCM. Cardiovascular events that required hospitalization increased when there was a family history (71.4% vs. 25.0%), which indicates the need to obtain a family medical history in the case of pregnancy with HCM.[6]

Pregnancy increases the circulating blood volume and cardiac output because of increases in the ventricular rate and stroke volume, while the peripheral vascular resistance decreases. The circulating blood volume increases more rapidly after 20 gestational weeks and reaches a plateau at 32 gestational weeks of 40–45% of the nonpregnant volume.[10,11] In HCM, the preload increase, afterload decrease and increase in cardiac contraction are precipitating factors because the ventricular blood volume decreases and left ventricular outflow obstruction deteriorates. The influence of these pregnancy-related physiologic changes on the circulation in HCM is not well understood. Therefore, in this the first study of this condition in Japan, we investigated the cardiovascular events that occurred during pregnancy with HCM.

Methods

We examined the outcomes of 27 pregnancies (23 women with HCM) between 1995 and 2013 at the Department of Perinatology, National Cerebral and Cardiovascular Center, Japan. HCM was diagnosed using the definition and type classification of cardiomyopathy published by the World Health Organization/International Society and Federation of Cardiology Joint Committee in 1995, by a cardiovascular physician based on medical history, physical findings, ECG, chest X-rays, an echocardiogram, and Doppler ultrasound. Radionuclide scans, computed tomography, magnetic resonance imaging, a cardiac catheter test, coronary arteriography, myocardial biopsy, and genetic diagnosis were performed when necessary. HCM was subcategorized into hypertrophic nonobstructive cardiomyopathy (HNCM), hypertrophic obstructive cardiomyopathy (HOCM), and dilated phase of HCM (D-HCM) with systolic dysfunction such as left ventricular ejection fraction (LVEF) <50%.

Information on maternal background was collected, including age, parity, complications, medications before pregnancy, NYHA class before pregnancy, family history of HCM, echocardiographic parameters; maximum wall thickness, LVEF, left atrial diameter in endsystole (LADs), mitral regurgitation (MR), and the pressure gradient (PG) of the left ventricular outflow tract obstruction (LVOTO), CARPREG score[12] and ZAHARA score[13] were retrospectively calculated. The CARPREG score is a contemporary assessment of maternal and neonatal risks

Case no.	LVEF <50%	Family history	LADs >50 mm	MR ≥moderate	LVOTO >50 mmHg	Maximau wall thickness >30 mm	ZAHARA score	CARPREG score
1	–	–	–	–	–	–	0	0
2	–	–	–	–	–	–	1.5	0
3	–	+	–	–	–	–	0	0
4	–	+	–	–	–	–	0	0
5	–	+	–	–	–	–	0	0
6	–	+	–	–	–	–	0	0
7	–	–	–	–	–	–	1.5	0
8	–	–	–	–	–	–	3	1
9	–	–	–	–	–	–	0	0
10	–	+	+	+	+	+	4.75	1
11	–	–	–	–	–	–	0	0
12	–	–	–	–	–	–	3	1
13	–	–	–	–	–	–	3	2
14	–	–	–	–	–	–	0	0
15	–	+	–	–	–	–	0	0
16	–	+	–	–	–	–	0	0
17	–	–	–	–	–	–	3	1
18	–	+	–	–	–	–	0	0
19	–	–	–	–	–	–	3	1
20	–	–	–	–	–	–	0	0
21	–	–	–	–	–	–	0	0
22	–	–	–	–	–	–	0	0
23	–	–	–	–	–	–	0	0
24	–	–	–	–	–	–	1.5	1
25	–	–	–	–	–	–	3	1.5
26	–	–	–	–	–	–	1.5	1
27	–	–	–	–	–	–	3	1.5

associated with pregnancy in women with heart disease who are receiving comprehensive prenatal care. Frequency of maternal primary cardiac events, as predicted by the risk index and observed in the derivation and validation groups, is expressed as a function of the number of cardiac predictors or points. The ZAHARA score is a modified risk score for cardiac complications during completed pregnancies in women with congenital heart disease.

Maternal and neonatal outcomes were examined, including cardiovascular events, NYHA class during pregnancy, NYHA class postpartum, gestational age, delivery mode, indication for cesarian section, birth weight, pH of the umbilical artery, and Apgar score at 5 min. Cardiovascular events were defined as new onset or worsening of arrhythmia, heart failure, endocarditis, or thromboembolic events that required medication, hospitalization, or termination of pregnancy. The gestational week of the occurrence of all cardiovascular events was recorded. Cardiovascular events were also classified as those related to arrhythmia or other than arrhythmia. The type of arrhythmia, gestational week of occurrence, and the detection method were recorded for each cardiovascular event related to arrhythmia.

Statistical Analysis

Univariate analysis by chi-squared test and the Cochran-Armitage trend test was used for statistical analysis. $P<0.05$ was considered significant.

Results

Maternal Background

Maternal background data for the 27 pregnancies (23 women) with HCM are shown in Table 1. The median age was 32 years (21–39 years). The mother was nulliparous in 17 pregnancies (63%) and multiparous in 10 (48%). Cases 3 and 4, 12 and 13, 15 and 16, and 21 and 22 relate to the same woman in each pair of cases (4 women). There were 17 women with HNCM, 6 with HOCM, and none D-HCM. One woman was complicated with tetralogy of Fallot after repair. Other maternal complicating diseases were idiopathic thrombocytopenic purpura and Bartter syndrome in 1 woman each. The medications administered before pregnancy were verapamil in 3 women, diltiazem in 2 women, β-blocker in 6 women, mexiletine in 2 women and angiotensin-converting enzyme inhibitor in 1 woman. A family history of HCM was identified in 6 women (26%). The NYHA class before pregnancy was I in all except 1 woman in class II (case 10) and that woman had LADs >50 mm, moderate MR, and a PG of LVOTO >50 mmHg. Among the other women with HOCM, the PG of LVOTO before pregnancy or in early pregnancy was between 15 and 35 mmHg. Therefore, all of the patients, except for the woman in case 10, were in good general condition.

Pregnancy Outcomes

Maternal and neonatal outcomes for the 27 pregnancies (23 women) with HCM are shown in Tables 2,3. A total of 17 cardiovascular events occurred in 13 pregnancies (48%), includ-

Table 2. Outcomes of the 27 Pregnancies in 23 Women With HCM

Case no.	Cardiovascular event	Gestational week	NYHA class (pregnancy)	NYHA class (postpartum)	Gestational age (weeks)	Delivery mode	Indication of CS	Birth weight (g)	UA pH	APS (5 min)
1	No		1	1	36	VD		2,690	7.36	9
2	Yes	30	1	1	38	VD		2,822	7.33	9
3	Yes	30	1	1	38	CS	NRFS	2,286	7.28	9
4	Yes	13	1	1	37	CS	Previous CS	2,940	7.35	9
5	Yes	32	1	1	40	VD		3,190	7.32	9
6	No		1	1	38	VD		2,724	7.25	9
7	Yes	32	1	1	36	VD		2,426	7.31	9
8	Yes	12, 31	1	1	31	CS	Heart	1,512	7.25	6
9	No		1	1	40	VD		3,016	7.32	9
10	Yes	13, 27	3	2	27	CS	Heart	850	7.29	5
11	Yes	32, 36	3	1	36	CS	Heart	2,250	7.21	7
12	Yes	13, 29	1	1	29	CS	Heart	1,013	7.33	3
13	Yes	Postpartum	1	1	37	CS	Previous CS	2,070	7.35	8
14	No		1	1	37	VD		2,326	7.24	10
15	No		1	1	39	VD		2,562	7.25	9
16	No		1	1	38	VD		3,124	7.38	10
17	Yes	31	1	1	39	VD		2,250	7.306	9
18	No		1	1	37	VD		2,874	7.201	8
19	Yes	Postpartum	1	1	37	VD		2,533	7.26	9
20	No		1	1	37	VD		3,364	7.35	9
21	No		1	1	38	VD		2,858	7.27	8
22	No		1	1	40	VD		3,054	7.29	8
23	No		1	1	39	CS	Previous CS	3,008	7.28	9
24	No		1	1	34	VD		2,434	7.35	10
25	Yes	24	1	1	39	CS	CPD	3,994	7.34	9
26	No		1	1	37	VD		2,674	7.29	10
27	Yes	Postpartum	1	1	38	VD		2,646	7.27	8

CPD, cephalopelvic disproportion; CS, cesarean section; NRFS, non-reassuring fetal status; UA, umbilical artery; VD, vaginal delivery. Other abbreviations as in Table 1.

Table 3. Summary of Maternal Outcomes of 27 Pregnancies With HCM

Cardiovascular event	14/27 (52%)
Termination of pregnancy because of cardiovascular event	4/27 (15%)
Worsening of NYHA class during pregnancy	2/27 (7%)
Preterm birth	7/27 (26%)

Abbreviations as in Table 1.

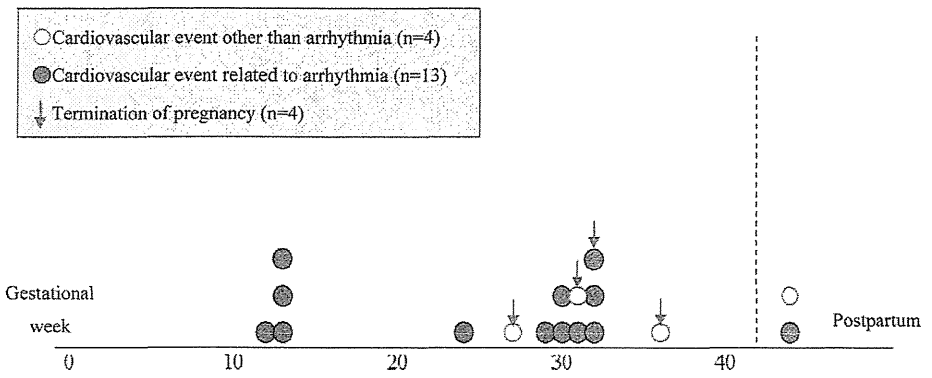

Figure. Cardiovascular events and gestational weeks (n=18).

Table 4. Arrhythmic Cardiovascular Events (n=13) During 27 Pregnancies in 23 Women With HCM

Case no.	Type of arrhythmia	Period (week)	Correspondence	Antiarrhythmic drug	Dose (mg)
2	PVC	30	Antiarrhythmic drug started	Metoprolol	30
3	NSVT	30	Antiarrhythmic drug started	Metoprolol	60
4	NSVT	13	Antiarrhythmic drug started	Propranolol	60
5	NSVT	32	Antiarrhythmic drug started	Metoprolol	40
7	PVC, PAC	32	Antiarrhythmic drug started	Metoprolol	40
8	PVC	12	Antiarrhythmic drug started	Bisoprolol	2.5
10	PVC	13	Antiarrhythmic drug started	Propranolol	60
11	NSVT	32	Antiarrhythmic drug started	Atenolol	25
12	NSVT	13	Antiarrhythmic drug increased	Metoprolol	40→60
12	NSVT	29	CS	–	–
17	PVC	31	Antiarrhythmic drug started	Propranolol	20
25	VT	24	Antiarrhythmic drug started	Carvedilol	10
27	PVC	Postpartum	Antiarrhythmic drug increased	Propranolol	60→80

NSVT, nonsustained ventricular tachycardia (VT); PAC, premature atrial contraction; PVC, premature ventricular contraction. Other abbreviations as in Tables 1,2.

ing 13 events (76%) related to arrhythmia (Table 4). Arrhythmia was the most common cardiovascular event. The cardiovascular events occurred in the early stage of pregnancy at approximately 30 gestational weeks, or postpartum (Figure). The events related to arrhythmia mainly occurred in the early stage of pregnancy or at approximately 30 gestational weeks. A total of 4 pregnancies were terminated because of a cardiovascular event (cases 8, 10, 11, 12). In case 8, the pregnancy was terminated at 31 gestational weeks because the mother was developing pulmonary hypertension and the PG of LVOTO had increased rapidly (peak PG 57 mmHg). Postpartum, the PG returned to the pre-pregnancy value. In cases 10 and 11, pregnancy was terminated at 27 and 36 gestational weeks, respectively, because in both cases there was an increased PG of LVOTO resulting from increased preload, and the mothers developed lung edema. After termination, the lung edema improved in both cases. In case 12, the pregnancy was terminated because nonsustained ventricular tachycardia (NSVT) could not be controlled with drug therapy. Thus, 3 of the 4 pregnancies (75%) were terminated because of a cardiovascular event in the mother who had started or increased her dose of antiarrhythmic drugs.

Premature delivery occurred in 7 of the 27 pregnancies (26%) because of cardiovascular events in 4 cases (57%) and obstetric complications (threatened premature labor) in 3 cases.

When comparing the pregnancies complicated by cardiovascular events with those unaffected by such events, the NYHA class before pregnancy, and echocardiographic parameters (LVEF, LADs, MR, LVOTO, maximum wall thickness) could not be analyzed because of the small number of positive findings. HOCM or family history of HCM were not risk factors (P=0.22, P=0.90). In the current study, medication in the pre-pregnancy period and CARPREG or ZAHARA score ≥1 or more were identified as risk factors of cardiac events during pregnancy or postpartum (Table 5).

Table 5. Relation of Cardiovascular Event and ZAHARA/CARPERG Score and Pre-Pregnancy Medication in 23 Women With HCM

	Cardiovascular events	P value
HCM		
No	8/20 (40%)	NS
Yes	5/7 (71%)	
D-HCM		
No	13/27 (48%)	NS
Yes	0/0 (0%)	
Medication (pre-pregnancy)		
No	5/18 (28%)	<0.05
Yes	8/9 (88%)	
NYHA class (pre-pregnancy)		
1	12/26 (46%)	NS
≥2	1/1 (100%)	
LVEF <50%		
No	13/27 (48%)	NS
Yes	0/0 (0%)	
Family history		
No	9/19 (33%)	NS
Yes	4/8 (50%)	
LADs >50 mm		
No	12/26 (46%)	NS
Yes	1/1 (100%)	
MR ≥ moderate		
No	12/26 (46%)	NS
Yes	1/1 (100%)	
LVOTO >50 mmHg		
No	12/26 (46%)	NS
Yes	1/1 (100%)	
Maximum wall thickness >30 mm		
No	12/26 (46%)	NS
Yes	1/1 (100%)	
High CARPREG score		
0	6/17 (35%)	<0.05
≥1	7/10 (70%)	
High ZAHARA score		
0	4/15 (26%)	<0.05
≥1	9/12 (75%)	

NS, not significant. Other abbreviations as in Table 1.

Discussion

A cardiovascular event related to HCM occurred in 13 of 27 pregnancies. Cardiovascular events showed 3 peak times of occurrence: early pregnancy, approximately 30 gestational weeks, and postpartum. In previous similar studies,[3–8] women who were symptomatic before pregnancy and who had a family history were at risk of cardiovascular events during their

pregnancies. Our new findings are that risk factors of cardiovascular events were medication before pregnancy and higher CARPREG or ZAHARA score.

The frequency of cardiovascular events (48%) is similar to the 28–73% reported in previous studies.[3–9] Collectively the findings show there is a high frequency of cardiovascular events in pregnancy for women with HCM. In the present study, 13 of the 18 events were related to arrhythmia, indicating that many of the cardiovascular events in pregnancy with HCM involve arrhythmia. Mostly, it was ventricular arrhythmias, including premature ventricular contraction and NSVT, and in some cases they were not controllable by medication, which is unusual. These findings indicate the importance of recognizing arrhythmia as a probable cardiovascular event in a pregnant woman with HCM.

Cardiovascular events occurred most frequently at approximately 30 gestational weeks. The increase in the circulating blood volume at 32 gestational weeks reaches 40–45% of the nonpregnant level, and it is notable that the most frequent period of cardiovascular events coincided approximately with the period of peak circulating blood volume during pregnancy. In 3 of the 4 pregnancies terminated because of a cardiovascular event, the event occurred during this period, which suggests that such cases require strict management and medication in the early stage of pregnancy.

Medication before pregnancy and higher CARPREG or ZAHARA score were risk factors for experiencing a cardiovascular event during pregnancy. However, further accumulation of cases and a study of multiple factors are required. These additional factors should include the general condition of the HCM patient, which appeared to influence the outcome in this study, and the observations from previous studies, which include an excellent perinatal prognosis in patients who are asymptomatic before pregnancy,[5] family history,[8] the apparent lack of influence of pregnancy on the natural course of HCM, and the tendency for a good prognosis when no symptoms are present before pregnancy.[7] Consideration of the timing of cardiovascular events may also be included in this analysis, given our finding of a high frequency of cardiovascular events in the early stage of pregnancy, at approximately 30 gestational weeks, and postpartum. Medication in the pre-pregnancy period, and CARPREG or ZAHARA score ≥1 were identified as risk factors of cardiac events during pregnancy or postpartum. However, this study was a retrospective analysis with the limitations of a small number of patients and the rarity of the condition.

Conclusions

If a pregnant woman with HCM has such factors as medication in the pre-pregnancy period or CARPREG or ZAHARA score ≥1, careful observation for cardiovascular events is required, especially at approximately 12 and 30 weeks' gestation and also postpartum.

References

1. Miura K, Nakagawa H, Morikawa Y, Nishikawa T, Yamamoto E, Ohki H, et al. Epidemiology of idiopathic cardiomyopathy in Japan: Results from a nationwide survey. *Heart* 2002; **87:** 126–130.
2. Matsumori A, Furukawa Y, Hasegawa K, Sato Y, Nakagawa H, Morikawa Y, et al. Epidemiology and clinical characteristics of cardiomyopathies in Japan: Results from nationwide surveys. *Circ J* 2002; **66:** 323–336.
3. Turner GM, Oakley CM, Dixon HG. Management of pregnancy complicated by hypertrophic obstructive cardiomyopathy. *BMJ* 1968; **4:** 281–284.
4. Oakley GD, McGarry K, Limb DG, Oakley CM. Management of pregnancy in patients with hypertrophic cardiomyopathy. *BMJ* 1979; **1:** 1749–1750.
5. Autore C, Conte MR, Piccininno M, Bernabò P, Bonfiglio G, Bruzzi P, et al. Risk associated with pregnancy in hypertrophic cardiomyopathy. *J Am Coll Cardiol* 2002; **40:** 1864–1869.
6. Avila WS, Rossi EG, Ramires JA, Grinberg M, Bortolotto MR, Zugaib M, et al. Pregnancy in patients with heart disease: Experience with 1,000 cases. *Clin Cardiol* 2003; **26:** 135–142.
7. Thaman R, Varnava A, Hamid MS, Firoozi S, Sachdev B, Condon M, et al. Pregnancy-related complications in women with hypertrophic cardiomyopathy. *Heart* 2003; **89:** 752–756.
8. Avila WS, Amaral FM, Ramires JA, Rossi EG, Grinberg M, Bortolotto MR, et al. Influence of pregnancy on clinical course and fetal outcome of women with hypertrophic cardiomyopathy. *Arq Bras Cardiol* 2007; **88:** 480–485.
9. Siu SC, Sermer M, Colman JM, Alvarez AN, Mercier LA, Morton BC, et al. Prospective multicenter study of pregnancy outcomes in women with heart disease. *Circulation* 2001; **104:** 515–521.
10. Katsuragi S. Maternal outcome in pregnancy complicated with pulmonary arterial hypertension. *Circ J* 2012; **76:** 2249–2254.
11. Kamiya CA. Outcome of pregnancy and effects on the right heart in women with repaired tetralogy of Fallot. *Circ J* 2012; **76:** 957–963.
12. Siu SC, Sermer M, Colman JM. Prospective multicenter study of pregnancy outcomes in women with heart disease. *Circulation* 2001; **104:** 515–521.
13. Drenthen W, Boersma E, Balci A. Predictors of pregnancy complications in women with congenital heart disease. *Eur Heart J* 2010; **31:** 2124–2132.

Eur J Pediatr (2014) 173:1537–1540
DOI 10.1007/s00431-013-2104-6

CASE REPORT

Refractory pulmonary hypertension following extremely preterm birth: paradoxical improvement in oxygenation after atrial septostomy

Hiroki Saitsu · Osuke Iwata · Junichiro Okada · Akiko Hirose · Hiroshi Kanda · Toyojiro Matsuishi · Kenji Suda · Yasuki Maeno

Received: 25 April 2013 / Revised: 9 July 2013 / Accepted: 12 July 2013 / Published online: 3 August 2013

Abbreviations

AcT/ET	Acceleration time relative to ejection time
BAS	Balloon atrial septostomy
CLD	Chronic lung disease
FiO_2	Inspired fractional oxygen concentration
PAH	Pulmonary arterial hypertension

Introduction

Despite newly introduced therapeutic regimens, outcomes of idiopathic pulmonary arterial hypertension (PAH) remain poor [5]. Recently, PAH associated with chronic lung disease (CLD) has been increasingly recognised, whose outcome is also unfavourable with a mortality rate of >30 % within 6 months of age [3, 7]. For idiopathic PAH refractory to pharmacological options, balloon atrial septostomy (BAS) has been proposed as a palliative intervention, aiming to prevent complete collapse of the circulation by releasing right atrial pressure through the atrial shunt [9]. However, little is known about the efficacy of BAS for PAH secondary to other clinical conditions. We present here a case of an extremely low-birth-weight infant who developed severe PAH associated with trisomy 21 and CLD, and whose PAH was permanently controlled by BAS with paradoxically improved arterial oxygenation.

H. Saitsu · O. Iwata · J. Okada · A. Hirose · H. Kanda · T. Matsuishi · K. Suda · Y. Maeno
Department of Paediatrics and Child Health, Kurume University School of Medicine, 67 Asahimachi, Kurume, Fukuoka, Japan

O. Iwata (✉)
Centre for Developmental and Cognitive Neuroscience, Kurume University School of Medicine, Kurume, Fukuoka 830-0011, Japan
e-mail: o.iwata@ucl.ac.uk

Case presentation

A 41-year-old woman, gravida 4, para 2, developed premature rupture of the membranes at 23 weeks and 5 days gestation; ritodrine hydrochloride, prophylactic antibiotics, and betamethasone were administered. A male neonate (birth weight, 725 g; Apgar scores, 5 and 7 at 1 and 5 min, respectively) was delivered by emergency caesarean section because of non-reassuring foetal status. The neonate required intubation for surfactant replacement therapy and mechanical ventilation; inspired fractional oxygen concentration (FiO_2) was swiftly reduced to <30 %. On echocardiography performed on day 0, transient pulmonary arterial hypertension (PAH) with the right-to-left dominant ductus arteriosus and trivial tricuspid valve regurgitation (TR) were observed, which resolved by day 2 when closure of the ductus arteriosus was confirmed; no cardiac defect was identified. Based on facial features characteristic of Down syndrome, a blood G banding test was performed, which confirmed trisomy 21. Mild leucopoenia and the presence of erythroblasts (with otherwise normal differential counts) were identified on a haemogram at birth, which resolved within the first week.

Chorioamnionitis was not observed on a placental histopathological examination; umbilical blood IgM was <5 mg/dL. However, the infant developed severe bronchopulmonary dysplasia (Fig. 1). Echocardiography on day 54 suggested mild PAH with shortened (28/155=0.18) acceleration time relative to the ejection time (AcT/ET) in the right ventricular outflow tract and a slightly increased (2.4 m/s) TR peak velocity. Although the infant was transiently extubated using nasal continuous positive airway pressure, abrupt and prolonged reductions in arterial oxygen saturation (SpO_2) <80 % with a poor response to 100 % oxygen (desaturation spells) were frequently observed corresponding to the patient's spontaneous activity. Despite treatments for CLD and

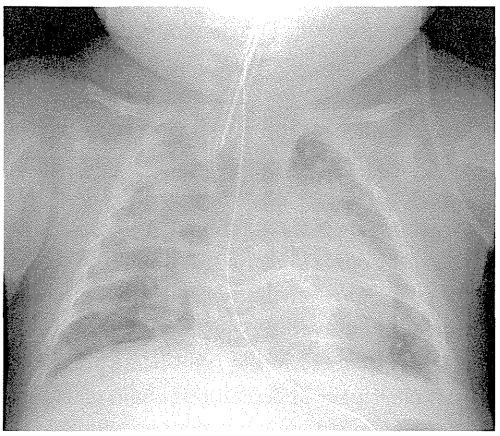

Fig. 1 Chest radiograph of the patient. Chest radiograph on day 79 showing severe lung injury with a mixed pattern of emphysema and atelectasis

PAH, including dexamethasone (started at 0.3 mg/kg/day followed by gradual weaning over 9 days), beraprost (2 μg/kg), sildenafil (3 mg/kg), and bosentan (4 mg/kg) (Fig. 2), the severity and frequency of the spells increased. Because severe spells were easily induced by spontaneous activity, continuous intravenous infusions of morphine (6 μg/kg/h) and rocuronium (4 μg/kg/min) were commenced. Following unsuccessful attempts to wean off deep sedation, tracheostomy was performed on day 282.

Echocardiography on day 376 showed consistently short AcT/ET (0.21) and a marked increase in the TR peak velocity (5.1 m/s), the latter of which was suggestive of super-systemic right ventricular pressure >100-mmHg. A foramen ovale with a diameter of 2-mm was observed, the flow of which was only visible during a desaturation spell in the direction of right to left. We concluded that an extension of conventional treatment was unlikely to improve the condition, but may merely prolong life.

After we obtained parental informed consent, on day 449, BAS was performed under general anaesthesia as a palliative treatment. Prior to BAS, the mean femoral and pulmonary artery pressures were 82 and 49-mmHg, respectively, and the left atrial pressure was 13-mmHg, giving a calculated pulmonary vascular resistance of 3.8 Wood units·m^2. With FiO_2 of 100 %, pulmonary vascular resistance decreased to 2.8 Wood

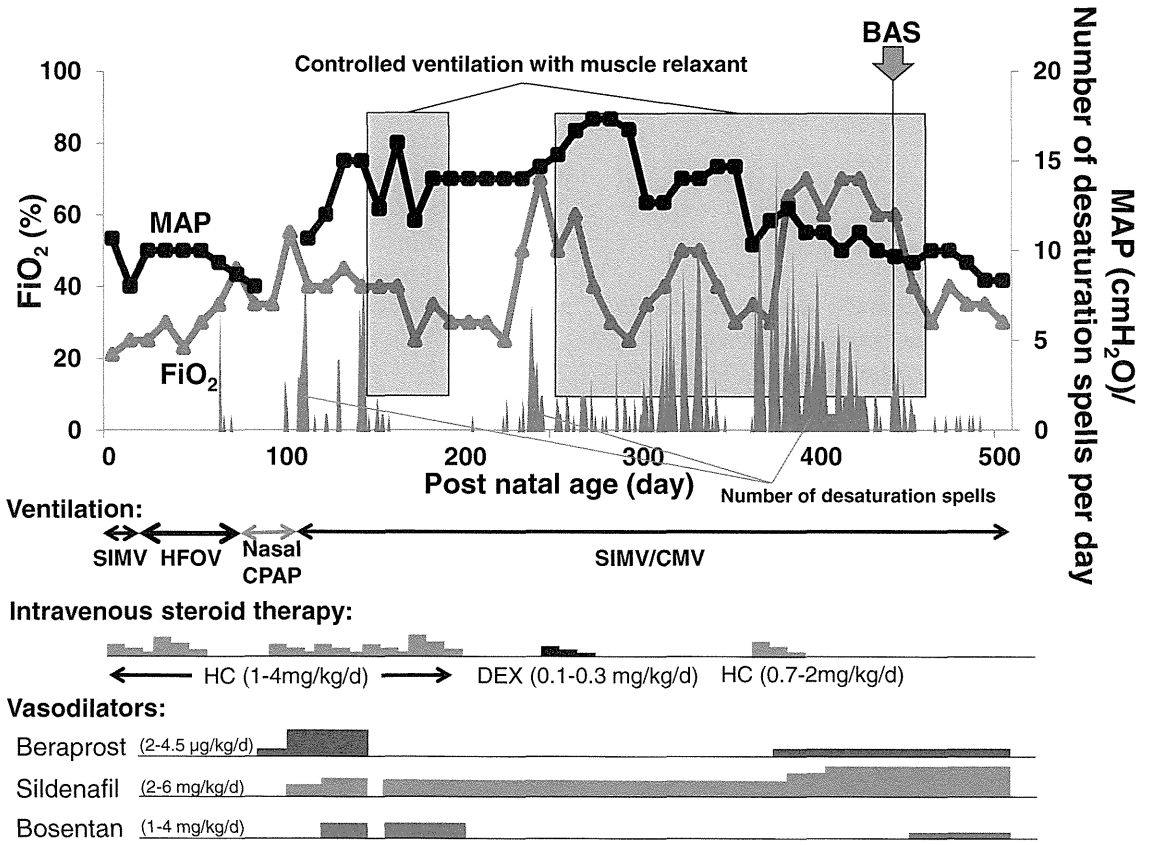

Fig. 2 Treatments and clinical variables over time. The patient required prolonged mechanical ventilation under deep sedation despite a range of medications shown in the lower panel. Minimum FiO_2 values required to maintain arterial haemoglobin oxygen saturation (SpO_2) >80 % at rest are shown. Desaturation spells are defined as SpO_2 <80 % longer than 10 min despite the use of 100 % oxygen (counted only once per hour when prolonged or repetitive). *CMV* conventional mechanical ventilation, *CPAP* continuous positive airway pressure, *DEX* dexamethasone, *FiO_2* fractional inspired oxygen, *HC* hydrocortisone, *HFOV* high-frequency oscillatory ventilation, *MAP* mean airway pressure, *SIMV* synchronised intermittent mandatory mechanical ventilation

units·m^2 with mean femoral and pulmonary artery pressures of 88 and 55-mmHg, respectively. When FiO_2 was reduced from 100 to 21 %, SpO_2 dropped from 93 to 80 %, and the mean pulmonary artery pressure increased to 118-mmHg, which was equivalent with the mean femoral artery pressure of 117-mmHg. BAS was then performed as a static balloon method using a 4×15-mm cutting balloon (Flextome®, Boston Scientific, Natick, MA, USA) with 10 atmospheres, followed by additional dilation using a 7×20-mm balloon (Ohicho II®, Kaneka Medical Products, Tokyo, Japan) with 18 atmospheres, which created a 5-mm hole at the atrial septum.

After BAS, bosentan (1 mg/kg) was re-prescribed to control the transient deterioration of PAH associated with the intervention. Enalapril (0.05 mg/kg) was used to reduce the elevated left atrial pressure (13-mmHg). Since the day of BAS, few spells of SpO_2 reduction were observed. Unexpectedly, baseline FiO_2 required to maintain SpO_2 >85 % was lowered to less than 30 % shortly after the intervention, resulting in successful weaning from deep sedation on day 465 and from mechanical ventilation on day 534. On day 595, echocardiography confirmed left to right atrial shunt and normalised AcT/ET of 0.35 (TR was too trivial to give its peak velocity), suggesting significant amelioration of PAH. The patient was discharged home on day 614 with home oxygen therapy. Severe desaturation spells were not observed until the latest visit of the patient to the follow-up clinic at 40 months old.

Discussion

We experienced a case of an extremely preterm infant with CLD and trisomy 21, who suffered from frequent episodes of severe PAH spells refractory to pharmacological treatments. However, unexpectedly, BAS permanently resolved these spells and improved baseline oxygenation, due possibly to different pathological mechanism from idiopathic PAH, where BAS is used for palliation.

CLD is a common complication of preterm birth [13]. As in our case, CLD is often accompanied by PAH [1, 3, 11]. Therapeutic regimens have been recently developed for PAH, such as prostaglandin I_2, phosphodiesterase type 5 inhibitor, and endothelin receptor inhibitor [4–6, 9–11, 14]. For PAH refractory to these options, BAS has been proposed as a palliative therapy. Micheletti et al. reported that BAS may improve right-heart function of patients with idiopathic PAH [9]. However, mortality associated with BAS (death <1 month after the procedure) is reported to be 5 to 16 % [8], highlighting the potential risk of the procedure for patients with severe PAH; it is suggested that the indication for BAS may be restricted to refractory cases with recurrent syncope and severe right-heart failure, and to patients awaiting pulmonary transplantation [9]. In our case, PAH was refractory to a range of pharmacological treatments, and the patient required 15 months of intensive mechanical ventilation under deep sedation. Together with objective data supporting the overall benefit of diverting right atrial blood to the systemic circulation, we concluded that our patient had an indication for BAS.

Based on a previous report, we initially expected that SpO_2 may be reduced by approximately 7 % after BAS because of mixing of the venous blood to the systemic flow [9]. However, BAS unexpectedly improved oxygenation during rest. In our case, a paroxysmal increase in the right-heart pressure might have been the main trigger of deleterious chain reactions, such as the release of endogenous excitatory catecholamines, inflammatory chemokines, and lung trauma due to high ventilator settings and high oxygen concentrations. Recent studies highlighted pivotal roles of inflammatory cytokines and growth factors in the progress of vascular remodelling and subsequent PAH [2]; improved control of right-heart pressure might help reduce the activation of the damaging cascade and assist the repair process of the pulmonary arterial endothelium. Future studies need to address the specific mechanism of PAH associated with underlying clinical conditions, such as CLD and Down syndrome [6, 12]. Although the benefit of BAS was obvious in our case, the indication of this procedure to young patients needs further investigation. As in patients with atrial septal defect, artificially created atrial septal communication may increase pulmonary flow. In the current case, inhaled nitric oxide was not considered because, at the time, only industrial gas was available, the use of which was strictly limited to perioperative or acute-phase patients; efficacy of inhaled nitric oxide and other pharmacological options also needs to be assessed in future studies.

Conclusions

Our findings highlighted that BAS may induce permanent remission of PAH with specific backgrounds, such as Down syndrome and CLD. However, given the risk associated with catheter interventions for patients with severe complications, the current indications of BAS for PAH should be strictly limited. Further prospective studies need to address the benefit, risk, and exact indication of this intervention for PAH secondary to Down syndrome, CLD, and other diseases.

Acknowledgments The authors are grateful to Dr. Yoshinori Kudo and Dr. Motofumi Iemura for their clinical input, to the nurses of the neonatal intensive care unit of Kurume University Hospital for their outstanding care, and to the patient and his family for agreeing to present the data. Dr. Iwata is funded by the Japan Science and Technology Agency and the Ministry of Education, Culture, Sports, Science and Technology (Grant-in-Aid for Scientific Research C24591533 and Grant-in-Aid for Scientific Research in Innovative Areas B01-24119004, Constructive Developmental Science), which did not have any role in the writing of the report and in the decision to submit the report for publication.

Conflict of interest The authors declare no competing conflict of interest.

References

1. Deakins KM (2009) Bronchopulmonary dysplasia. Respir Care 54:1252–1262
2. El Chami H, Hassoun PM (2012) Immune and inflammatory mechanisms in pulmonary arterial hypertension. Prog Cardiovasc Dis 55:218–228
3. Farquhar M, Fitzgerald DA (2010) Pulmonary hypertension in chronic neonatal lung disease. Paediatr Respir Rev 11:149–153
4. Farrow KN, Steinhorn RH (2011) Phosphodiesterases: emerging therapeutic targets for neonatal pulmonary hypertension. Handb Exp Pharmacol 204:251–277
5. Hawkins A, Tulloh R (2009) Treatment of pediatric pulmonary hypertension. Vasc Health Risk Manag 5:509–524
6. Hawkins A, Langton-Hewer S, Henderson J, Tulloh RM (2011) Management of pulmonary hypertension in Down syndrome. Eur J Pediatr 170:915–912
7. Khemani E, McElhinney DB, Rhein L, Andrade O, Lacro RV, Thomas KC, Mullen MP (2007) Pulmonary artery hypertension in formerly premature infants with bronchopulmonary dysplasia: clinical features and outcomes in the surfactant era. Pediatrics 120:1260–1269
8. Klepetko W, Mayer E, Sandoval J, Trulock EP, Vachiery JL, Dartevelle P, Pepke-Zaba J, Jamieson SW, Lang I, Corris P (2004) Interventional and surgical modalities of treatment for pulmonary arterial hypertension. J Am Coll Cardiol 43:73S–80S
9. Micheletti A, Hislop AA, Lammers A, Bonhoeffer P, Derrick G, Rees P, Haworth SG (2006) Role of atrial septostomy in the treatment of children with pulmonary arterial hypertension. Heart 92:969–972
10. Rashid A, Ivy D (2005) Severe paediatric pulmonary hypertension: new management strategies. Arch Dis Child 90:92–98
11. Rosenzweig EB, Widlitz AC, Barst RJ (2004) Pulmonary arterial hypertension in children. Pediatr Pulmonol 38:2–22
12. Seki M, Kato T, Masutani S, Matsunaga T, Senzaki H (2009) Pulmonary arterial hypertension associated with gastroesophageal reflux in a 2-month-old boy with Down syndrome. Circ J 73:2352–2354
13. Smith VC, Zupancic JA, McCormick MC, Croen LA, Greene J, Escobar GJ, Richardson DK (2005) Trends in severe bronchopulmonary dysplasia rates between 1994 and 2002. J Pediatr 146:469–473
14. Valerio CJ, Coghlan JG (2009) Bosentan in the treatment of pulmonary arterial hypertension with the focus on the mildly symptomatic patient. Vasc Health Risk Manag 5:607–619

Brief Report—Endocrine Research

Noninvasive Surrogate Markers for Plasma Cortisol in Newborn Infants: Utility of Urine and Saliva Samples and Caution for Venipuncture Blood Samples

Hisayoshi Okamura, Masahiro Kinoshita, Hiroki Saitsu, Hiroshi Kanda, Sachiko Iwata, Yasuki Maeno, Toyojiro Matsuishi, and Osuke Iwata

Centre for Developmental and Cognitive Neuroscience (H.O., S.I., T.M., O.I.) and Department of Paediatrics and Child Health (M.K., H.S., H.K., S.I., Y.M., T.M., O.I.), Kurume University School of Medicine, Kurume, Fukuoka 830-0011, Japan; and Division of Pediatrics (H.S., O.I.), St Mary's Hospital, Kurume, Fukuoka, 830-8543 Japan

Context: Hypothalamus-pituitary-adrenal function is associated with important physiological/pathological events in neonates. Plasma/serum cortisol levels have been used to assess hypothalamus-pituitary-adrenal function. Several noninvasive surrogate markers have been used without sufficient validation.

Objective: The objectives of the study were to investigate whether plasma cortisol levels are correlated with those in saliva and urine and whether these correlations are affected by procedural pain at blood sampling.

Design, Setting, and Patients: Fifty neonates were recruited from a tertiary neonatal intensive care unit. Saliva and urine samples were collected shortly before routine clinical blood sampling.

Main Outcome Measures: Cortisol levels were compared between plasma and noninvasive samples using a linear regression analysis for the entire study population and groups, whose blood was obtained via indwelling arterial catheters (group A) or by venipuncture (group V). Predictive values of salivary/urinary cortisol for low plasma cortisol levels less than 2.0 μg/dL were evaluated by receiver-operating characteristic analysis.

Results: Plasma cortisol showed linear correlations with salivary and urinary cortisol for the entire study population and patients in group A (all $P < .0002$) but not in group V. Areas under the curves of salivary and urinary cortisol to predict low plasma cortisol levels were 0.87 (95% confidence interval 0.78–0.97) and 0.84 (95% confidence interval 0.74–0.95), respectively.

Conclusions: Cortisol levels from saliva or urine samples were shown to be useful surrogate markers for plasma cortisol levels in neonates. Caution is required in interpreting the findings of plasma cortisol levels in young patients when blood samples are obtained by venipuncture because procedural pain may induce alteration of cortisol levels. (*J Clin Endocrinol Metab* 99: E2020–E2024, 2014)

A range of physiological and pathological conditions in neonates depend on the hypothalamus-pituitary-adrenal (HPA) function. Increased fetal plasma cortisol induces lung maturation and causes adaptation to extrauterine stress (1). Cortisol administration ameliorates chronic lung disease and pressor-resistant hypotension in immature infants (2, 3). Further understanding of the HPA axis regulation may considerably improve the treatment of high-risk neonates. HPA function has primarily been assessed using plasma cortisol (4); for young patients, blood sampling under low-stress conditions is recommended to minimize pain-induced cortisol elevation (5).

ISSN Print 0021-972X ISSN Online 1945-7197
Printed in U.S.A.

Received April 9, 2014. Accepted July 21, 2014.
First Published Online July 31, 2014

Abbreviations: AUC, areas under ROC curves; CI, confidence interval; HPA, hypothalamus-pituitary-adrenal; ROC, receiver-operating characteristic.

Salivary cortisol has been used as a surrogate marker of plasma cortisol (6), which has additionally been validated in preterm-born infants during the CRH stimulation test (7). This needs to be confirmed under physiological conditions, in which the dynamic range of cortisol is approximately one third of that with CRH stimuli (8). Urinary cortisol has also been used to estimate plasma cortisol levels (9), the relevance of which also needs to be confirmed in neonates.

This study aimed to investigate whether salivary and urinary cortisol reflects the plasma cortisol levels under physiological conditions and whether the relationship in the cortisol levels between noninvasive samples and blood samples differs according to the blood sampling procedure.

Materials and Methods

This study was conducted with the approval of the Ethics Committee of Kurume University School of Medicine. Written informed consent was obtained from a parent of each neonate.

Study population

Fifty neonates were recruited from a tertiary neonatal intensive care center (Kurume University Hospital, Kurume, Fukuoka, Japan) between December 2011 and August 2013. Neonates who were in a life-threatening condition, too immature (body weight < 1000 g and/or corrected age < 32 wk), anemic (blood hemoglobin < 12 g/dL and/or on erythropoietin therapy), and/or on cortisol supplementation were not included. Clinical variables of participants, such as gestational age, birth weight, antenatal/postnatal corticosteroid administration, and delivery mode, were collected from the patient's record.

Sample collection

Sample collection was performed on the day when regular blood sampling was scheduled. In our unit, regular blood sampling is performed at 10:00 AM to allow sufficient time after the latest 3-hourly feeding at 7:00 AM. A rectangular cotton pad (5 × 6 × 0.3 cm) was inserted within the diaper at 9:00 AM with maximum caution not to stimulate the neonates. The diaper was checked every 30 minutes until urination was confirmed. The cotton pad was then secured within a sealable plastic bag. Saliva samples were collected using an absorbent swab stick (Sorbette; Salimetrics) as previously reported (8). A sufficient amount of saliva sample was assumed when saliva oozed out of the swab end when pressed to the inner wall of the conical tube. A blood sample was then obtained either by venipuncture or via an indwelling arterial catheter when applicable. In our unit, arterial catheters are inserted for neonates with a gestational age of less than 26 weeks, unstable cardiopulmonary conditions, or scheduled surgical operations. Blood sampling for a cortisol assay (0.6 mL) was performed immediately after routine blood sampling (0.2–1.2 mL). We aimed to collect saliva and blood samples within 10–15 minutes after the urine was collected. All samples were then centrifuged at 4000 rpm for 10 minutes and were kept at −80°C until assayed. Repeated sampling from the same newborn was limited only to those who were healthy with a hemoglobin level of greater than 15 mg/dL, with an interval between samplings of more than 2 weeks.

Cortisol assay

Minimum sample volumes required for the assay were 50 μL (urine), 25 μL (saliva), and 100 μL (plasma). Salivary cortisol levels were determined by enzyme immunoassay (high sensitivity salivary cortisol ELISA kit; Salimetrics). The limit of detection of this assay in our laboratory was 0.19 nmol/L, and the intra- and interassay coefficients of variation were 5.43% and 6.41%, respectively. Plasma and urinary cortisol were measured using a cortisol enzyme immunoassay kit (Cayman Chemical). The limit of detection of this assay in our laboratory was 1.24 nmol/L, and the intra- and interassay coefficients of variation were 7.61% and 8.22%, respectively.

Data analysis

To minimize the bias caused by the sample collection, analysis was performed only for patients with a complete data set of plasma, salivary, and urinary cortisol. The correlations among plasma, salivary, and urinary cortisol of the overall samples were assessed using the Pearson's correlation coefficient. To examine the effect of procedural pain on the relationships between salivary/urinary cortisol and plasma cortisol, correlation analyses were repeated for groups of newborns, whose blood was collected by venipuncture (group V) or via an indwelling arterial catheter (group A) (Bonferroni corrected for two groups). Baseline characteristics and cortisol levels of neonates were compared between groups using the Student's t test and χ^2 test where applicable (Bonferroni corrected for 10 clinical variables and three sample types). Finally, threshold values of salivary/urinary cortisol to predict low plasma cortisol levels of less than 55 nmol/L [2.0 μg/dL: determined based on the reference range in term born neonates, below which dynamic tests are recommended (5)] were assessed using the receiver-operating characteristic (ROC) analysis.

Results

Data profile

Overall study population

Two neonates were excluded because of insufficient saliva volume. Two other newborns were studied twice each with intervals of 21 and 16 days. Eventually, 50 complete sample sets from 48 neonates were analyzed. Indications for hospitalization for these neonates were low birth weight (n = 33), congenital blood disease (n = 4), maternal gestational diabetes mellitus (n = 4), congenital heart disease (n = 3), neonatal encephalopathy (n = 2), infection (n = 2), maternal hyperthyroidism (n = 1), and suspicion of metabolic disease (n = 1) (see Table 1 for other profiles). Urinary cortisol levels were the highest, which were followed by those in the plasma and salivary samples (all difference, $P < .0003$) (Table 1).

Table 1. Background Clinical Variables

	Overall Subjects (n = 50)	Group A (n = 25)	Group V (n = 25)
Female	23 (46%)	7 (28%)	16 (64%)
Birth weight, g[a]	1897 (833)	1545 (800)	2248 (721)
Gestational age, wk[b]	34.1 (4.7)	31.0 (4.7)	36.8 (2.6)
Postnatal age, d	7.1 (7.8)	7.2 (7.5)	7.1 (8.3)
Apgar score at 1 min[b]	6.9 (2.3)	5.4 (2.4)	7.7 (1.4)
Apgar score at 5 min[a]	8.2 (1.5)	7.6 (1.8)	8.8 (0.8)
Caesarean section	25 (50%)	15 (60%)	10 (40%)
Antenatal corticosteroid	16 (32%)	12 (48%)	4 (16%)
Postnatal corticosteroid[c]	9 (18%)	9 (36%)	0 (0%)
Days on respiratory support[d,e]	18.4 (28.7)	32.1 (33.2)	2 (3.9)
Cortisol level, nmol/L			
Plasma	64.7 (54.3)	76.1 (61.1)	53.3 (47.4)
Saliva[e]	15.2 (10.3)	20.7 (13.7)	9.7 (6.9)
Urine	113.5 (74.5)	134.7 (90.2)	92.3 (58.8)

Values are shown as the mean (SD) or the number of cases (percentage).

[a] $P < .05$; [b] $P < .001$; and [c] $P < .01$ from between-group comparisons corrected for multiple comparisons over 10 clinical background variables using the Bonferroni correction.

[d] Including nasal continuous positive airway pressure, mechanical ventilation, and oxygen supplemental care.

[e] $P < .005$ from between-group comparisons corrected for multiple comparisons over 10 clinical background variables and three sample types using the Bonferroni correction.

Group characteristics

Neonates in group A had a smaller birth weight, gestational age, and Apgar scores at 1 and 5 minutes compared with those in group V ($P < .05$, $P < .001$, $P < .001$, and $P < .05$, respectively) (Table 1). More newborns in group A were administered postnatal corticosteroids and had more days of respiratory support compared with those in group V ($P < .01$ and $P < .005$, respectively); other variables were similar between the two groups.

Relationships between plasma, salivary and urinary cortisol

Overall study population

Plasma cortisol levels were positively correlated with salivary ($r = 0.71$, $P < .0002$) and urinary ($r = 0.61$, $P < .0002$) cortisol levels. Salivary and urinary cortisol levels also showed a linear correlation between each other ($r = 0.68$, $P < .0002$) (Figure 1, A–C).

Group analysis

Plasma cortisol levels were positively correlated with salivary and urinary cortisol levels in group A ($r = 0.82$ and 0.67, respectively; both $P < .0002$), but not in group V ($r = 0.43$ and 0.42, respectively) (Figure 1, A–C).

Receiver-operating characteristic analysis

For the entire study population, mean areas under ROC curves (AUCs) to predict low plasma cortisol levels were 0.87 [95% confidence interval (CI) 0.78–0.97] and 0.84 (95% CI 0.74–0.95) for salivary and urinary cortisol, respectively (Figure 1D). The optimal sensitivity and specificity were obtained with cutoff values of 12.7nmol/L (sensitivity 0.76, specificity 0.83) for salivary cortisol and 109.0nmol/L (sensitivity 0.76, specificity 0.83) for urinary cortisol. In group A, the AUCs improved to 0.94 (95% CI 0.85–1.00) for salivary cortisol and 0.87 (95% CI 0.73–1.00) for urinary cortisol, whereas in group V, the AUC decreased to 0.77 (95% CI 0.58–0.95) for salivary cortisol and 0.80 (95% CI 0.62–0.98) for urinary cortisol (Figure 1, E and F).

Discussion

Salivary and urinary cortisol are well-established surrogate markers for plasma cortisol in adults and adolescents (6, 9). Consistent to a previous observation in immature neonates under CRH stimulation (7), we observed a robust linear correlation between salivary and plasma cortisol levels under physiological conditions, thus expanding the utility of salivary cortisol for virtually all neonates. In a previous study, sufficient saliva samples for the assay were collected only in 46% of infants (10), which contrasts with a high success rate of 96% in our study. This might be associated with our protocol to check the approximate sample volume during collection using a simple maneuver (see *Sample collection* for details). However, saliva collection may still be burdensome for longitudinal studies over days. In our study, we additionally assessed urinary cortisol levels, which showed a marked correlation with plasma cortisol values. Urine collection using cotton pads or disposable diapers is simple and noninvasive (11). The use of urinary samples provides additional information, such as total cortisol secretion per day when serially collected (12). However, this technique might not be optimal when the exact timing of changes in cortisol levels is important because urinary cortisol is likely to reflect plasma cortisol levels when urine is extracted from blood, rather than when urine is collected (13). Urinary cortisol levels may also be affected by the amount of urination; correction of urinary cortisol for creatinine may

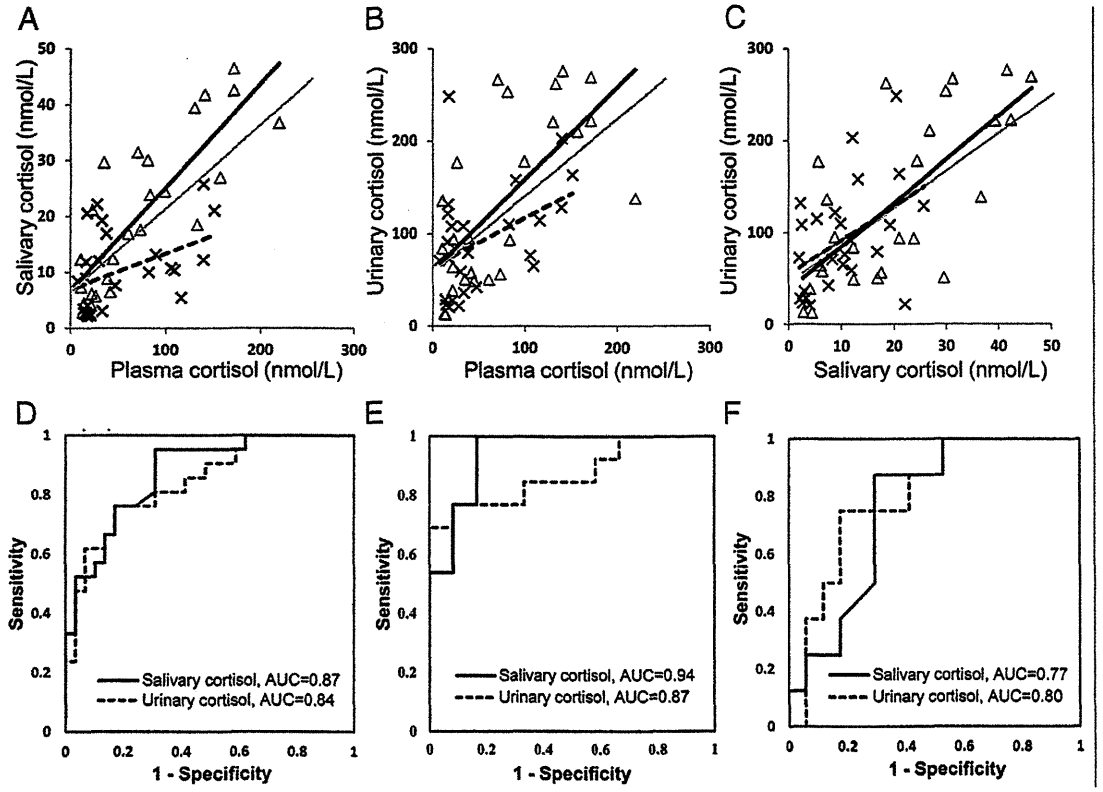

Figure 1. Relationships among plasma, salivary, and urinary cortisol levels. A–C, Scatter plot showing the association among plasma (arterial or venous), salivary, and urinary cortisol levels. Plasma cortisol levels showed strong linear correlations with salivary (A) and urinary (B) cortisol levels for the overall population ($r = 0.71$ and $r = 0.61$, respectively; both $P < .0002$) and group A neonates (blood samples obtained via an indwelling arterial catheter; $r = 0.82$ and $r = 0.67$, respectively; both $P < .0002$) but not for group V neonates (blood samples obtained by venipuncture; $r = 0.43$ and $r = 0.42$, respectively). D and E, ROC curves of salivary and urinary cortisol to predict low plasma cortisol levels less than 55 nmol/L. D, Overall study population. AUCs were 0.87 (95% CI 0.78–0.97) for salivary cortisol and 0.84 (95% CI 0.74–0.95) for urinary cortisol. Optimal cutoff values for salivary and urinary cortisol were 12.7nmol/L (sensitivity 0.76, specificity 0.83) and 109.0nmol/L (sensitivity 0.76, specificity 0.83), respectively. E, Neonates in group A. AUCs improved to 0.94 (95% CI 0.85–1.00) for salivary cortisol and 0.87 (95% CI 0.73–1.00) for urinary cortisol. Optimal cutoff values for salivary and urinary cortisol were 14.6 nmol/L (sensitivity 1.00, specificity 0.83) and 136.9 nmol/L (sensitivity 0.77, specificity 0.92), respectively. F, Neonates in group V. The AUC was 0.77 (95% CI 0.58–0.95) for salivary cortisol and 0.80 (95% CI 0.62–0.98) for urinary cortisol; optimal cutoff values for salivary and urinary cortisol were 9.5nmol/L (sensitivity 0.88, specificity 0.71) and 109.0nmol/L (sensitivity 0.75, specificity 0.82), respectively. △, Group A; ×, group V. Gray lines, Regression for the whole population; bold lines, group A; broken, group V.

improve the reliability of this technique (14). These limitations need to be considered for the use of urine cortisol in routine clinical practice, in which considerable intra- and interpatient variability is observed in the amount and interval of urination.

Our secondary analysis showed that robust linear correlations between plasma and salivary/urinary cortisol levels were observed only for neonates, whose blood was sampled via an indwelling arterial catheter. Painful procedures induce sympathetic stimulation and subsequent increase in circulating adrenocorticotropic hormone within less than a few minutes (15). Our findings indicate that the assessment of plasma cortisol using venipuncture blood might not be reliable, at least for neonates. Studies in the 1980s through the 1990s observed an overt diurnal rhythm in plasma cortisol only after a few months of birth (16, 17). Other studies failed to reduce a cortisol response to painful procedures by analgesia (18). Although these findings were attributed to immature HPA functioning, recent studies, which used salivary cortisol markers, showed that diurnal cortisol rhythms were present, even shortly after birth (8); other studies have shown that nonnutritive sucking, maternal skin-to-skin contact, and gentle sensory stimuli reduced excessive adrenal responses to procedural pain in infants (19, 20). Taken together, blood sampling via venipuncture may induce a prompt elevation in cortisol according to the response of each neonate, rendering these invasive samples unsuitable for assessment of adrenal function at a young age. Current understanding of the developing HPA axis may carefully be reappraised when knowledge is obtained from studies that use invasive blood samples.

Limitations of the study

Because of ethical reasons, arterial and venous blood samples were not obtained from the same patient. There-

fore, caution for cortisol levels of venipuncture blood is suggested only from indirect observations. Although we expected pain-induced elevation of plasma cortisol in some group V patients, cortisol levels were relatively higher in group A. This finding could be explained by the greater severity of condition and postnatal corticosteroid administration for newborns, who required an arterial catheter. Finally, our study population was recruited from a neonatal intensive care center, whose clinical background and environmental factors were different from their healthy peers.

Conclusions

Salivary and urinary cortisol levels can be used for both snapshot and serial monitoring of plasma cortisol levels in neonates. Blood samples obtained by painful procedures may not be suitable for the assessment of resting plasma cortisol levels. Specific conditions, in which cortisol levels of venipuncture blood become unreliable, need to be identified.

Acknowledgments

We thank the patients who participated in the study and their parents for their cooperation; the nurses of the neonatal intensive care unit of Kurume University Hospital for their support; Dr Yuko Araki for her guidance in the statistical analysis; Drs Mitsuaki Unno, Naoko Hara, Junichiro Okada, Tadashi Hisano, and Akiko Hirose for their contribution to the data collection; and Ms Chihoko Urata, Chiho Yoshii, and Chiaki Ueno for their consistent support and encouragement.

Address all correspondence and requests for reprints to: Osuke Iwata, MD, Centre for Developmental and Cognitive Neuroscience, Department of Pediatrics and Child Health, Kurume University School of Medicine, 67 Asahimachi, Kurume, Fukuoka 830–0011, Japan. E-mail: o.iwata@ucl.ac.uk.

This work was supported by the Japan Society for the Promotion of Science and The Ministry of Education, Culture, Sports, Science, and Technology (Grant-in-Aid for Scientific Research C24591533, Grants-in-Aid for Young Scientists B 24730613 and B01–24119004, Constructive Developmental Science, Innovative Areas).

Disclosure Summary: The authors have nothing to declare.

References

1. **Bolt RJ, van Weissenbruch MM, Lafeber HN, Delemarre-van de Waal HA.** Development of the hypothalamic-pituitary-adrenal axis in the fetus and preterm infant. *J Pediatr Endocrinol Metab.* 2002; 15:759–769.
2. **Doyle LW, Davis PG, Morley CJ, McPhee A, Carlin JB.** Low-dose dexamethasone facilitates extubation among chronically ventilator-dependent infants: a multicenter, international, randomized, controlled trial. *Pediatrics.* 2006;117:75–83.
3. **Okada J, Iwata S, Hirose A, et al.** Levothyroxine replacement therapy and refractory hypotension out of transitional period in preterm infants. *Clin Endocrinol (Oxf).* 2011;74:354–364.
4. **Kudielka BM, Buske-Kirschbaum A, Hellhammer DH, Kirschbaum C.** HPA axis responses to laboratory psychosocial stress in healthy elderly adults, younger adults, and children: impact of age and gender. *Psychoneuroendocrinology.* 2004;29:83–98.
5. **Donohoue P, Wolfsdorf J, Pierrat A.** Diagnosis of adrenal insufficiency in children. In: Post TW, ed. UpToDate. Waltham, MA: UpToDate; 2014.
6. **Voegtline KM, Granger DA.** Dispatches from the interface of salivary bioscience and neonatal research. *Front Endocrinol.* 2014;5: 25.
7. **Matsukura T, Kawai M, Marumo C, et al.** Diagnostic value of salivary cortisol in the CRH stimulation test in premature infants. *J Clin Endocrinol Metab.* 2012;97:890–896.
8. **Iwata O, Okamura H, Saitsu H, et al.** Diurnal cortisol changes in newborn infants suggesting entrainment of peripheral circadian clock in utero and at birth. *J Clin Endocrinol Metab.* 2013;98:E25–E32.
9. **Morineau G, Boudi A, Barka A, et al.** Radioimmunoassay of cortisone in serum, urine, and saliva to assess the status of the cortisol-cortisone shuttle. *Clin Chem.* 1997;43:1397–1407.
10. **Herrington CJ, Olomu IN, Geller SM.** Salivary cortisol as indicators of pain in preterm infants: a pilot study. *Clin Nurs Res.* 2004;13: 53–68.
11. **Alam MT, Coulter JB, Pacheco J, et al.** Comparison of urine contamination rates using three different methods of collection: clean-catch, cotton wool pad and urine bag. *Ann Trop Paediatr.* 2005; 25:29–34.
12. **Siamopoulou-Mavridou A, Mavridis AK, Vizandiadis A, Harsoulis P.** Free urinary cortisol immunoreactive levels in premature and full term infants. *Acta Paediatr Scand.* 1986;75:919–922.
13. **Lasley BL, Mobed K, Gold EB.** The use of urinary hormonal assessments in human studies. *Ann NY Acad Sci.* 1994;709:299–311.
14. **Heckmann M, Hartmann MF, Kampschulte B, et al.** Assessing cortisol production in preterm infants: do not dispose of the nappies. *Pediatr Res.* 2005;57:412–418.
15. **Meeran K, Hattersley A, Mould G, Bloom SR.** Venepuncture causes rapid rise in plasma ACTH. *Br J Clin Pract.* 1993;47:246–247.
16. **Mantagos S, Moustogiannis A, Vagenakis AG.** Diurnal variation of plasma cortisol levels in infancy. *J Pediatr Endocrinol Metab.* 1998; 11:549–553.
17. **Vermes I, Dohanics J, Toth G, Pongracz J.** Maturation of the circadian rhythm of the adrenocortical functions in human neonates and infants. *Horm Res.* 1980;12:237–244.
18. **Stang HJ, Gunnar MR, Snellman L, Condon LM, Kestenbaum R.** Local anesthesia for neonatal circumcision. Effects on distress and cortisol response. *JAMA.* 1988;259:1507–1511.
19. **Cong X, Ludington-Hoe SM, Walsh S.** Randomized crossover trial of kangaroo care to reduce biobehavioral pain responses in preterm infants: a pilot study. *Biol Res Nurs.* 2011;13:204–216.
20. **South MM, Strauss RA, South AP, Boggess JF, Thorp JM.** The use of non-nutritive sucking to decrease the physiologic pain response during neonatal circumcision: a randomized controlled trial. *Am J Obstet Gynecol.* 2005;193:537–542; discussion 542–533.

Research Article

Statistics in Medicine

Received 9 May 2013, Accepted 2 March 2014 Published online 27 March 2014 in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.6154

Sample size determination in group-sequential clinical trials with two co-primary endpoints

Koko Asakura,[a,b] Toshimitsu Hamasaki,[a,b*†] Tomoyuki Sugimoto,[c] Kenichi Hayashi,[a] Scott R. Evans[d] and Takashi Sozu[e]

We discuss sample size determination in group-sequential designs with two endpoints as co-primary. We derive the power and sample size within two decision-making frameworks. One is to claim the test intervention's benefit relative to control when superiority is achieved for the two endpoints *at the same interim timepoint* of the trial. The other is when superiority is achieved for the two endpoints *at any interim timepoint, not necessarily simultaneously*. We evaluate the behaviors of sample size and power with varying design elements and provide a real example to illustrate the proposed sample size methods. In addition, we discuss sample size recalculation based on observed data and evaluate the impact on the power and Type I error rate. Copyright © 2014 John Wiley & Sons, Ltd.

Keywords: average sample number; conditional power; Cui–Hung–Wang statistics; co-primary endpoints; group-sequential methods; maximum sample size; sample size recalculation; Type I error

1. Introduction

Traditionally, in clinical trials, a single outcome is selected as a primary endpoint. This endpoint is then used as the basis for the trial design including sample size determination, interim monitoring, and final analyses. However, many recent clinical trials, especially in medical product development, have utilized more than one endpoint as *co-primary*. 'Co-primary' in this setting means that the trial is designed to evaluate if the new intervention is superior to the control on *all* endpoints, thus evaluating the intervention's multidimensional effects. Regulators have issued guidelines recommending co-primary endpoints in some disease areas. For example, the Committee for Medicinal Products for Human Use issued a guideline [1] recommending the use of cognitive, functional, and global endpoints to evaluate symptomatic improvement of dementia associated with Alzheimer's disease, indicating that primary endpoints should be stipulated reflecting the cognitive and functional disease aspects. Offen *et al.* [2] provides other examples with co-primary endpoints for regulatory purposes.

The resulting need for new approaches to the design and analysis of clinical trials with co-primary endpoints has been noted [2–4]. Utilizing multiple endpoints may provide the opportunity for characterizing intervention's multidimensional effects and also create challenges. Specifically controlling the Type I and Type II error rates when the multiple co-primary endpoints are potentially correlated is non-trivial. When designing the trial to evaluate the joint effects on *all* of the endpoints, no adjustment is needed to

[a]*Department of Biomedical Statistics, Osaka University Graduate School of Medicine, Osaka, Japan*
[b]*Office of Biostatistics and Data Management, Research and Development Initiative Center, National Cerebral and Cardiovascular Center, Osaka, Japan*
[c]*Department of Mathematical Sciences, Graduate School of Science and Technology, Hirosaki University, Aomori, Japan*
[d]*Department of Biostatistics and the Center for Biostatistics in AIDS Research, Harvard School of Public Health, Boston, MA, U.S.A.*
[e]*Department of Biostatistics, Kyoto University School of Public Health, Kyoto, Japan*
**Correspondence to: Toshimitsu Hamasaki, Office of Biostatistics and Data Management, Research & Development Initiative Center, National Cerebral and Cardiovascular Center, 5-7 Fujishiro-dai, Suita, Osaka 565-8565, Japan.*
[†]*E-mail: toshi.hamasaki@ncvc.go.jp*

control the Type I error rate. However, the Type II error rate increases as the number of endpoints to be evaluated increases. Thus, adjustments in design (i.e., sample size) are needed to maintain the overall power. Methods for clinical trials with co-primary endpoints have been discussed in fixed sample size designs by many authors [5–16]. Even if the correlation among the endpoints is incorporated into the sample size calculation, existing methods often result in large and impractical sample sizes as the testing procedure for co-primary endpoints is conservative. Chuang-Stein *et al.* [7] and Kordzakhia *et al.* [10] discuss the methods to adjust the significance levels that depend on the correlation among the endpoints in the fixed sample size designs. The methods may provide smaller sample sizes and also introduce the other challenges. For example, the sample size calculated to detect the joint effect may be smaller than the sample size calculated for each individual endpoint. The prespecified correlation incorporated into the significance level adjustment is usually unknown and may be incorrect. This calls into question whether or not the significance level should be updated based on the observed correlation.

In this paper, we extend previous work for the fixed sample size designs, considering sample size evaluation in the group-sequential setting with co-primary endpoints. As suggested by Hung and Wang [3], a group-sequential design may be a remedial but practical approach because it offers the possibility to stop a trial early when evidence is overwhelming and thus offers efficiency (i.e., potentially fewer patients than the fixed sample size designs). We discuss the case of two positively correlated continuous outcomes. We consider a two-arm parallel-group trial designed to evaluate if an experimental intervention is superior to a control. The paper is structured as follows. In Section 2, we describe the statistical setting, decision-making frameworks for rejecting the null hypothesis, and definitions of power. In Section 3, we evaluate the behaviors of sample size and power with varying design elements and then provide a real example to illustrate the methods. In Section 4, we describe sample size recalculation and the resulting effect on power and Type I error rate. In Section 5, we summarize the findings and discuss the further developments.

2. Group-sequential designs with two co-primary endpoints

2.1. Statistical setting

Consider a randomized, group-sequential clinical trial of comparing the test intervention (T) with the control intervention (C). Two continuous outcomes are to be evaluated as co-primary endpoints. Suppose that a maximum of L analyses are planned, where the same number of analyses with the same information space are selected for both endpoints. Let n_l and $r_l n_l$ be the cumulative number of participants on the test and the control intervention groups at the lth analysis ($l = 1, \ldots, L$), respectively, where r_l is the sampling ratio. Hence, up to n_L and $r_L n_L$ participants are recruited and randomly assigned to the test and the control intervention groups, respectively. Then, there are n_L paired outcomes $(Y_{\mathrm{T}1i}, Y_{\mathrm{T}2i})$ $(i = 1, \ldots, n_L)$ for the test intervention group and $r_L n_L$ paired outcomes $(Y_{\mathrm{C}1j}, Y_{\mathrm{C}2j})$ $(j = 1, \ldots, r_L n_L)$ for the control intervention group. Assume that $(Y_{\mathrm{T}1i}, Y_{\mathrm{T}2i})$ and $(Y_{\mathrm{C}1j}, Y_{\mathrm{C}2j})$ are independently bivariate-normally distributed as $(Y_{\mathrm{T}1i}, Y_{\mathrm{T}2i}) \sim \mathrm{N}_2(\mu_{\mathrm{T}1}, \mu_{\mathrm{T}2}, \sigma^2_{\mathrm{T}1}, \sigma^2_{\mathrm{T}2}, \rho_{\mathrm{T}})$ and $(Y_{\mathrm{C}1j}, Y_{\mathrm{C}2j}) \sim \mathrm{N}_2\left(\mu_{\mathrm{C}1}, \mu_{\mathrm{C}2}, \sigma^2_{\mathrm{C}1}, \sigma^2_{\mathrm{C}2}, \rho_{\mathrm{C}}\right)$, respectively. For simplicity, the variances are assumed to be known and common, that is, $\sigma^2_{\mathrm{T}1} = \sigma^2_{\mathrm{C}1} = \sigma^2_1$ and $\sigma^2_{\mathrm{T}2} = \sigma^2_{\mathrm{C}2} = \sigma^2_2$. Note that the method can be applied to the case of unknown variances. For the fixed sample size designs, Sozu *et al.* [12] discuss a method for the unknown variance case and show that the calculated sample size is nearly equivalent to that for the known variance in the setting of a one-sided significance level $\alpha = 0.025$ and power $1 - \beta = 0.8$ or 0.9. By analogy from the fixed sample designs, there is no practical difference in the group-sequential setting, and the methodology for a known variance provides a reasonable approximation for the unknown variances case.

Let (δ_1, δ_2) denote the differences in the means for the test and the control intervention groups, respectively, where $\delta_k = \mu_{\mathrm{T}k} - \mu_{\mathrm{C}k} (k = 1, 2)$. Suppose that positive values of (δ_1, δ_2) represent the test intervention's benefit. We are interested in conducting a hypothesis test to evaluate if the intervention is superior to the control intervention, that is, the null hypothesis $\mathrm{H}_0 : \delta_1 \leqslant 0$ or $\delta_2 \leqslant 0$ versus the alternative hypothesis $\mathrm{H}_1 : \delta_1 > 0$ and $\delta_2 > 0$. Let (Z_{1l}, Z_{2l}) be the statistics for testing the hypotheses at the lth analysis, given by $Z_{kl} = (\bar{Y}_{\mathrm{T}kl} - \bar{Y}_{\mathrm{C}kl}) / (\sigma_k \sqrt{\kappa_l / n_l})$, where $\kappa_l = (1 + r_l)/r_l$, and $\bar{Y}_{\mathrm{T}kl}$ and $\bar{Y}_{\mathrm{C}kl}$ are the sample means given by $\bar{Y}_{\mathrm{T}kl} = n_l^{-1} \sum_{i=1}^{n_l} Y_{\mathrm{T}ki}$ and $\bar{Y}_{\mathrm{C}kl} = (r_l n_l)^{-1} \sum_{j=1}^{r_l n_l} Y_{\mathrm{C}kj}$. Z_{1l} and Z_{2l} are normally distributed as $\mathrm{N}(\sqrt{n_l/\kappa_l}\delta_1/\sigma_1, 1^2)$ and $\mathrm{N}(\sqrt{n_l/\kappa_l}\delta_2/\sigma_2, 1^2)$, respectively. Thus, (Z_{1l}, Z_{2l}) is bivariate-normally distributed with the correlation $(r_l \rho_{\mathrm{T}} + \rho_{\mathrm{C}})/(1 + r_l)$. Furthermore, the

joint distribution of $(Z_{11}, Z_{21}, \ldots, Z_{1L}, Z_{2L})$ is $2L$ multivariate normal with their correlations given by $\mathrm{corr}[Z_{kl}, Z_{k'l'}] = \sqrt{\kappa_l n_{l'} / \kappa_{l'} n_l}$ if $k = k'$; $\sqrt{\kappa_l n_{l'}}(r_l \rho_T + \rho_C) / \{\sqrt{\kappa_{l'} n_l}(1 + r_l)\}$ if $k \neq k'$.

2.2. *Decision-making framework, stopping rules, and power*

When evaluating the joint effects on both of the endpoints within the context of group-sequential designs, there are the two decision-making frameworks associated with hypothesis testing. One is to reject H_0 if and only if superiority is achieved for the two endpoints simultaneously (i.e., at the same interim timepoint of the trial) (DF-1). The other is to reject H_0 if superiority is achieved for the two endpoints at any interim timepoint (i.e., not necessarily simultaneously) (DF-2). We will discuss the two decision-making frameworks separately as the corresponding stopping rules and power definitions are unique.

DF-1: The DF-1 is relatively simple: if superiority is demonstrated on only one endpoint at an interim, then the trial continues and the hypothesis testing is repeated for both endpoints until the joint significance for the two endpoints is established simultaneously. The stopping rule for DF-1 is formally given as follows:

At the lth analysis $(l = 1, \ldots, L-1)$

If $Z_{1l} > c_{1l}$ and $Z_{2l} > c_{2l}$,	then reject H_0 and stop the trial,
otherwise,	continue to the $(l+1)$th analysis,

at the Lth analysis

if $Z_{1L} > c_{1L}$ and $Z_{2L} > c_{2L}$,	then reject H_0,
otherwise,	do not reject H_0,

where c_{1l} and c_{2l} are the critical values, which are constant and selected separately, using any group-sequential method such as the Lan–DeMets (LD) alpha-spending method [17] to control the overall Type I error rate of α, as if they were a single primary endpoint, ignoring the other co-primary endpoint. The testing procedure for co-primary endpoints is conservative. For example, if a zero correlation between the two endpoints is assumed and each endpoint is tested at the one-sided significance level of 2.5%, then the Type I error rate is 0.0625%. As shown in Section 4, the maximum Type I error rate associated with the rejection region of the null hypothesis increases as the correlation goes toward one, but it is not greater than the targeted significance level.

The power corresponding to DF-1 is

$$1 - \beta = \Pr\left[\bigcup_{l=1}^{L} \{A_{1l} \cap A_{2l}\} \middle| H_1\right], \tag{1}$$

where $A_{kl} = \{Z_{kl} > c_{kl}\} (k = 1, 2;\ l = 1, \ldots, L)$. The power (1) can be numerically assessed by using multivariate normal integrals. A detailed calculation is provided in Appendix A.1.

DF-2: DF-2 is more flexible than DF-1. If superiority is demonstrated on one endpoint at the interim, then the trial will continue, but subsequent hypothesis testing is repeatedly conducted only for the previously non-significant endpoint until superiority is demonstrated. The stopping rule for DF-2 is formally given as follows:

At the lth analysis $(l = 1, \ldots, L-1)$

If $Z_{1l} > c_{1l}$ and $Z_{2l'} > c_{2l'}$ for some $1 \leqslant l' \leqslant l$,	then reject H_0 and stop the trial,
if $Z_{2l} > c_{2l}$ and $Z_{1l'} > c_{1l'}$ for some $1 \leqslant l' \leqslant l$,	then reject H_0 and stop the trial,
otherwise,	continue to the $(l+1)$th analysis,

at the Lth analysis

if $Z_{1L} > c_{1L}$ and $Z_{2l'} > c_{2l'}$ for some $1 \leqslant l' \leqslant L$,	then reject H_0,
if $Z_{2L} > c_{2L}$ and $Z_{1l'} > c_{1l'}$ for some $1 \leqslant l' \leqslant L$,	then reject H_0,
otherwise,	do not reject H_0.

Therefore, following DF-2, the power is

$$1 - \beta = \Pr\left[\left\{\bigcup_{l=1}^{L} A_{1l}\right\} \cap \left\{\bigcup_{l=1}^{L} A_{2l}\right\} \middle| H_1\right]. \tag{2}$$