



Validation of Fetal Medicine Foundation charts for fetal growth in twins: nationwide Danish cohort study

S. E. KRISTENSEN^{1,2,3} , A. WRIGHT⁴, D. WRIGHT⁴, K. GADSBØLL^{1,2}, C. K. EKELEND^{1,2}, P. SANDAGER^{5,6,7}, F. S. JØRGENSEN^{2,8}, E. HOSETH⁹, L. SPERLING¹⁰, H. J. ZINGENBERG¹¹, K. SUNDBERG¹, A. MCLENNAN^{12,13} , K. H. NICOLAIDES¹⁴ and O. B. PETERSEN^{1,2}

¹Center for Fetal Medicine, Pregnancy and Ultrasound, Department of Gynecology, Fertility and Obstetrics, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark; ²Faculty of Health and Medical Sciences, Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark; ³Faculty of Health and Medical Sciences, Institute of Cellular and Molecular Medicine, University of Copenhagen, Copenhagen, Denmark; ⁴Institute of Health Research, University of Exeter, Exeter, UK; ⁵Department of Obstetrics and Gynecology, Center for Fetal Medicine, Aarhus University Hospital, Aarhus, Denmark; ⁶Center for Fetal Diagnostics, Aarhus University Hospital, Aarhus, Denmark; ⁷Department of Clinical Medicine, Aarhus University, Aarhus, Denmark; ⁸Department of Obstetrics and Gynecology, Fetal Medicine Unit, Copenhagen University Hospital – Hvidovre and Amager, Hvidovre, Denmark; ⁹Department of Obstetrics and Gynecology, Clinic of Ultrasound, Aalborg University Hospital, Aalborg, Denmark; ¹⁰Department of Obstetrics and Gynecology, Center for Ultrasound and Pregnancy, Odense University Hospital, Odense, Denmark; ¹¹Department of Obstetrics and Gynecology, Ultrasound in Pregnancy, Copenhagen University Hospital – Herlev and Gentofte, Herlev, Denmark; ¹²Sydney Ultrasound for Women, Chatswood, NSW, Australia; ¹³Discipline of Obstetrics, Gynaecology and Neonatology, The University of Sydney, Sydney, NSW, Australia; ¹⁴Harris Birthright Research Centre for Fetal Medicine, King's College Hospital, London, UK

KEYWORDS: dichorionic; external validation; fetal growth model; growth monitoring; monochorionic diamniotic; national cohort; singleton reference distribution; small-for-gestational age; twin pregnancy; twin reference distribution

CONTRIBUTION

What are the novel findings of this work?

This study used a nationwide Danish cohort of 8542 dichorionic and 1675 monochorionic diamniotic twin pregnancies to externally validate the Fetal Medicine Foundation (FMF) chorionicity-specific models for fetal growth in twin pregnancies compared with singleton percentiles. The dual-percentile scale may offer improved resolution compared with the conventional scale in g.

What are the clinical implications of this work?

The FMF chorionicity-specific models provide a way of assessing fetal growth in twins relative to singletons. Future research should seek to improve the accuracy, validity and clinical utility of these models for growth monitoring in twin pregnancy.

ABSTRACT

Objective To assess the validity of the Fetal Medicine Foundation (FMF) chorionicity-specific models for fetal growth in twin pregnancy.

Methods This was an external validation study of the FMF models using a nationwide Danish cohort of

twin pregnancies. The cohort included all dichorionic (DC) and monochorionic diamniotic (MCDA) twin pregnancies with an estimated delivery date between 2008 and 2018, which satisfied the following inclusion criteria: two live fetuses at the first-trimester ultrasound scan (11–14 weeks' gestation); biometric measurements available for the calculation of estimated fetal weight (EFW) using the Hadlock-3 formula; and delivery of two liveborn infants. Validation involved assessing the distributional properties of the models and estimating the mean EFW Z-score deviations. Additionally, the models were applied to pregnancies that delivered preterm and attended non-scheduled visits (complicated pregnancies).

Results Overall, 8542 DC and 1675 MCDA twin pregnancies met the inclusion criteria. In DC twins, 17 084 fetuses were evaluated at a total of 95 346 ultrasound scans, of which 44.5% were performed at scheduled visits in pregnancies carried to 37 + 0 weeks or later. The median number of growth scans per DC twin fetus from 20 + 0 weeks onwards was four. The model showed good agreement with the validation cohort for scheduled visits in DC twins delivered at 37 + 0 weeks or later (mean \pm SD EFW Z-score, -0.14 ± 1.05). In MCDA twins, 3350 fetuses underwent 31 632 eligible

Correspondence: Dr S. E. Kristensen, Center for Fetal Medicine and Ultrasound, Department of Gynecology, Fertility and Obstetrics, Copenhagen University Hospital, Rigshospitalet, Blegdamsvej 9, DK-2100 Copenhagen, Denmark (e-mail: steffen.ernesto.kristensen.01@regionh.dk)

Accepted: 26 September 2024

ultrasound scans, of which 59.5% were performed at scheduled visits in pregnancies carried to 36+0 weeks or later. The median number of growth scans per MCDA twin fetus from 16+0 weeks onwards was 10. The model showed favorable agreement with the validation cohort for scheduled visits in MCDA twins delivered at 36+0 weeks or later (mean \pm SD EFW Z-score, -0.09 ± 1.01). Non-scheduled visits and preterm delivery before 37+0 weeks for DC twins and before 36+0 weeks for MCDA twins corresponded with smaller weight estimates, which was consistent with the study's definition of complicated pregnancy.

Conclusions The FMF models provide a good fit for EFW measurements in our Danish national cohort of uncomplicated twin pregnancies assessed at routine scans. Therefore, the FMF models establish robust criteria for subsequent investigations and potential clinical applications. Future research should focus on exploring the consequences of clinical implementation, particularly regarding the identification of twins that are small-for-gestational age, as they are especially susceptible to adverse perinatal outcome. © 2024 The Author(s). *Ultrasound in Obstetrics & Gynecology* published by John Wiley & Sons Ltd on behalf of International Society of Ultrasound in Obstetrics and Gynecology.

INTRODUCTION

Twin pregnancies present unique challenges in obstetrics, necessitating tailored monitoring and antenatal care to ensure optimal outcomes for the mother and the twins. Accurate assessment of fetal growth is crucial in managing twin pregnancy, but it remains complex owing to inherent physiological variations and pathological conditions associated with multifetal pregnancy and chorionicity. Twin pregnancies account for 1.4% of all pregnancies in Europe¹. However, the proportion of twins requiring admission to the neonatal intensive care unit is much higher than that for singletons, partly owing to the higher proportion of infants being born preterm but also because of impaired fetal growth and the risks associated with low birth weight^{2–5}.

In conventional obstetric practice, the categorization of fetal development into small-for-gestational age (SGA), appropriate-for-gestational age (AGA) or large-for-gestational age was based on standard growth charts developed for singleton pregnancies. These growth charts serve as universal references, based on the assumption that a fetus has the same genetic growth potential regardless of the plurality of the pregnancy, thereby facilitating a uniform methodology for assessing fetal growth across all pregnancies and enabling comparisons with the general population^{6–9}.

The Fetal Medicine Foundation (FMF) has recently introduced a reference for twins on a singleton percentile scale¹⁰. This enables growth in twins to be assessed against both singleton and twin reference distributions (Figure 1). The additional information on fetal growth provided by

this reference compared with a twin-specific reference could offer a better understanding of the complexities associated with these high-risk pregnancies^{11–14}.

The objective of this study was to externally validate the FMF models for describing chorionicity-specific fetal growth in twin pregnancies relative to that in singleton pregnancies.

METHODS

Study population and design

This was an external validation study of the FMF models using the nationwide Danish cohort of twin pregnancies. The cohort included all dichorionic (DC) and monochorionic diamniotic (MCDA) twins with an estimated delivery date between 1 January 2008 and 31 December 2018 that satisfied the inclusion criteria listed below. The cohort was based on national highly detailed data from all local obstetric departments stored in the five regional Astraia servers¹⁵ combined with data obtained from the Danish Fetal Medicine Database over a span of 11 years, as described previously^{16,17}. In Denmark, the public health-care system provides ultrasound scans and first-trimester aneuploidy screening to all pregnant women between 11 and 14 weeks' gestation. The participation rate for this screening program exceeded 95% of all pregnant women during the study period¹⁸. Determination of chorionicity is performed for every twin pregnancy in the first trimester by evaluating the placentae and the intertwin membrane for the presence of the lambda sign or the T-sign^{19,20}. Additionally, obstetric departments across Denmark consistently comply with national guidelines for the management of twin pregnancies. These guidelines include a recommendation for scheduled visits at 4-weekly intervals for uncomplicated DC twins starting from 20 weeks, and at 2-weekly intervals for uncomplicated MCDA twins commencing from 16 weeks. Furthermore, the guidelines recommend delivery of DC twins at approximately 37–38 weeks and of MCDA twins at around 36–37 weeks.

Inclusion criteria for the study were a pregnancy with two live fetuses at the first-trimester ultrasound scan resulting in the birth of two liveborn infants. Exclusion criteria included terminated pregnancy, miscarriage before 24+0 weeks, stillbirth at or after 24+0 weeks, intrauterine demise of one fetus, pregnancy undergoing fetal reduction, missing outcome information or inadequate growth assessment to calculate estimated fetal weight (EFW) using the Hadlock-3 formula^{21,22}. Pregnancies affected by twin-specific complications such as selective fetal growth restriction, twin–twin transfusion syndrome or twin anemia polycythemia sequence were not specifically excluded, unless they did not fulfill the inclusion criteria. Cases in which chorionicity was uncertain and could not be determined through electronic medical records were excluded from further analysis. These exclusion criteria were chosen to align with those used in the development cohort¹⁰. Abnormal Doppler

measurements were not included in the development cohort, thus they were not included in this study.

The study received approval from the regional authority responsible for data security management (approval number: P-2020-381) and the Danish clinical quality registry (RKKP-case number: FØTO-2021-03-22). The Danish Patient Safety Authority authorized the retrieval of missing or supplementary data (case number: 31-1521-356). It should be noted that, following Danish legislation, ethical approval is not required for registry-based studies.

Outcome measures

The study defined uncomplicated DC and MCDA twin pregnancies as those seen for scheduled visits, resulting in two liveborn children delivered at or after 37 + 0 weeks in DC twin pregnancies and at or after 36 + 0 weeks in MCDA twin pregnancies. These cut-offs were determined based on international guidelines for the recommended timing of delivery in twin pregnancies^{23,24}. Scheduled visits were defined as those occurring at 4-weekly intervals in DC twins from 20 weeks' gestation and at 2-weekly intervals in MCDA twins from 16 weeks^{23,24}. Non-scheduled visits refer to those that took place between the scheduled visits. Pregnancies with non-scheduled visits and those that delivered preterm were considered as complicated.

EFW was calculated using the Hadlock-3 formula, using head circumference, abdominal circumference and femur length^{21,22}. EFW values were normalized as Z-scores (adjusted by gestational age) and calculated using the FMF reference for singletons⁶.

Statistical analysis

Chorionicity-specific fetal growth in twins described by the FMF models¹⁰, relative to singletons, is illustrated in Figure 1. The models were developed by fitting Bayesian hierarchical models using Markov chain Monte Carlo (MCMC) simulation methods²⁵. Moreover, the models employed a hierarchical Gaussian framework with three levels: (1) pregnancy; (2) fetus; and (3) visit. The models were fitted to singleton Z-scores for uncomplicated DC and MCDA twin pregnancies.

For the primary validation of the models, a subgroup of uncomplicated twin pregnancies in the validation cohort was used. For comparison, an assessment was also made of non-scheduled visits and pregnancies that delivered before 37 + 0 weeks in DC twins or before 36 + 0 weeks in MCDA twins. This assessment aimed to assess deviations from the FMF reference in non-scheduled growth assessments and in pregnancies that delivered preterm. Summary statistics and Gaussian probability plots of Z-scores

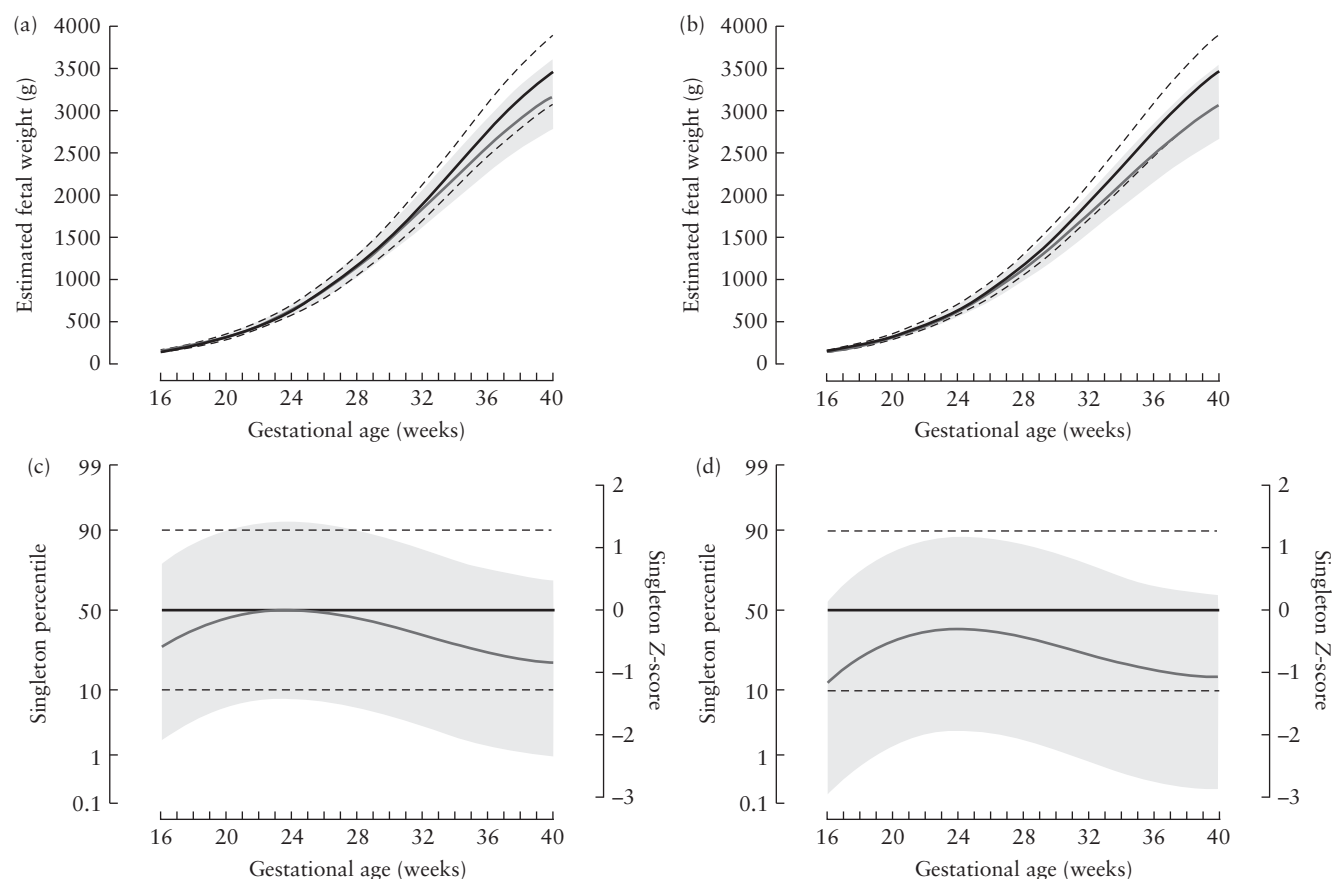


Figure 1 Fetal Medicine Foundation chorionicity-specific fetal growth charts for dichorionic (a,c) and monochorionic diamniotic (b,d) twins, in relation to singleton estimated fetal weights, percentiles and Z-scores, reproduced from Wright *et al.*¹⁰. Solid black line is 50th percentile and black dashed lines are 10th and 90th percentiles for singletons. Solid gray line is 50th percentile and shaded area represents range between 10th and 90th percentiles for twins.

for EFW in MCDA and DC twins were generated to assess the accuracy of the models. The Z-scores were distributed from the models for scheduled and non-scheduled visits according to term or preterm delivery and for each scheduled visit in both chorionicities. Data analysis was carried out using the statistical software RStudio (version 2023.06.0+421)²⁶.

RESULTS

In the cohort of 12 380 twin pregnancies, 10 077 (81.4%) were DC, of which 8542 (84.8%) met the inclusion criteria, and 2116 (17.1%) were MCDA, of which 1675 (79.2%) met the inclusion criteria²⁷. Reasons for exclusion are given in Figure 2.

Baseline characteristics

Maternal baseline characteristics and obstetric history are summarized in Table 1. Most women were in their late twenties or early thirties. There were few active smokers and approximately half of the women were nulliparous. The majority identified as Caucasian and had a pregestational body mass index within the normal range. Natural conception accounted for half of the

DC pregnancies compared with 85.0% of the MCDA pregnancies.

Validation of dichorionic twin model

The DC twin cohort comprised 17 084 fetuses evaluated at a total of 95 346 eligible ultrasound scans. Among these scans, 68.4% ($n = 65\,209$) were performed at scheduled visits and 44.5% ($n = 42\,453$) were performed at scheduled visits in pregnancies carried to 37 + 0 weeks or later. The median number of scans per fetus from 20 + 0 weeks onwards was four (interquartile range (IQR), 3–5).

Figure 3 and Table 2 show the distributional properties of EFW Z-scores derived from the model for DC twins. The model exhibited favorable alignment with the validation cohort for scheduled visits of DC twin pregnancies delivered at or after 37 + 0 weeks (mean \pm SD EFW Z-score, -0.14 ± 1.05). However, for non-scheduled visits of pregnancies delivered at or after 37 + 0 weeks, a larger proportion of Z-scores was observed in the lower range, indicating a left shift in the distribution (mean \pm SD EFW Z-score, -0.31 ± 1.17). Furthermore, pregnancies delivered before 37 + 0 weeks, both for scheduled visits (mean \pm SD EFW Z-score, -0.35 ± 1.31) and non-scheduled visits (mean \pm SD EFW Z-score, -0.69 ± 1.50), exhibited even smaller weight

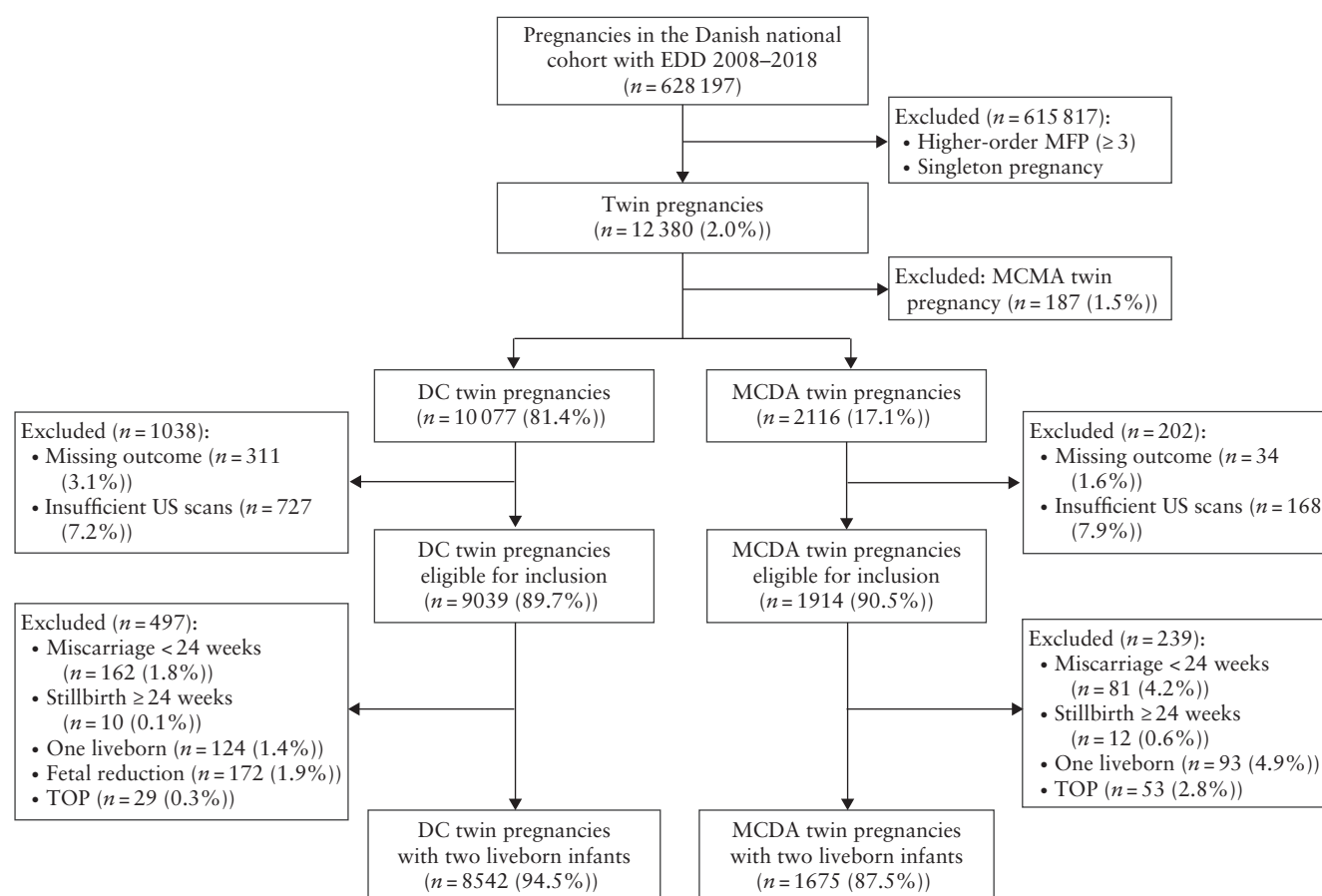


Figure 2 Flowchart summarizing inclusion of patients in study population. DC, dichorionic; EDD, estimated delivery date; MCDA, monochorionic diamniotic; MCMA, monochorionic monoamniotic; MFP, multifetal pregnancy; TOP, termination of pregnancy; US, ultrasound.

Table 1 Baseline maternal characteristics of twin pregnancies included in validation cohort

Characteristic	DC twin pregnancy (n = 8542)	MCDA twin pregnancy (n = 1675)
Fetuses (n)	17 084	3350
Maternal age (years)	32 (29–35)	30 (27–33)
Maternal height (cm)*	169 (164–173)	168 (164–172)
Maternal weight (kg)*	67 (60–78)	65 (59–76)
Prepregnancy BMI (kg/m ²)*	24 (21–27)	23 (21–27)
Conception		
Natural	4253/8353 (50.9)	1391/1637 (85.0)
Ovarian induction	430/8353 (5.1)	11/1637 (0.7)
Intrauterine insemination	914/8353 (10.9)	47/1637 (2.9)
In-vitro fertilization	2716/8353 (32.5)	185/1637 (11.3)
Egg donation	40/8353 (0.5)	3/1637 (0.2)
Nulliparous	2620/5553 (47.2)	592/1249 (47.4)
Current smoker	586/8343 (7.0)	125/1550 (8.1)
Caucasian ethnicity	7835/8235 (95.1)	1523/1637 (93.0)

Data are given as *n*, median (interquartile range) or *n/N* (%).

*Data missing for 199 dichorionic (DC) and 39 mono chorionic diamniotic (MCDA) twin pregnancies. BMI, body mass index.

estimates (Z-scores), corresponding to a further left shift in the Z-score distributions. These findings align with our definition of complicated DC pregnancy, namely one characterized by non-scheduled visits and delivery before 37 + 0 weeks. Figure S1 provides a visual representation of the distributional properties of Z-scores at each 4-weekly scheduled visit and the overall distribution for scheduled visits in pregnancies delivered at 37 + 0 weeks or later.

Validation of mono chorionic diamniotic twin model

The cohort of MCDA twins included 3350 fetuses assessed during 31 632 eligible ultrasound scans. Among these scans, 91.5% (*n* = 28 934) were performed at scheduled visits and 59.5% (*n* = 18 816) were performed at scheduled visits in pregnancies carried to 36 + 0 weeks or later. The median number of scans per fetus from 16 + 0 weeks onwards was 10 (IQR, 8–12).

Figure 4 and Table 2 display the distributional properties of EFW Z-scores in the validation cohort using the model for MCDA twins. The model exhibited favorable alignment with the validation cohort for scheduled visits of MCDA twin pregnancies delivered at 36 + 0 weeks or later (mean ± SD EFW Z-score, -0.09 ± 1.01). However, for non-scheduled visits of pregnancies delivered at

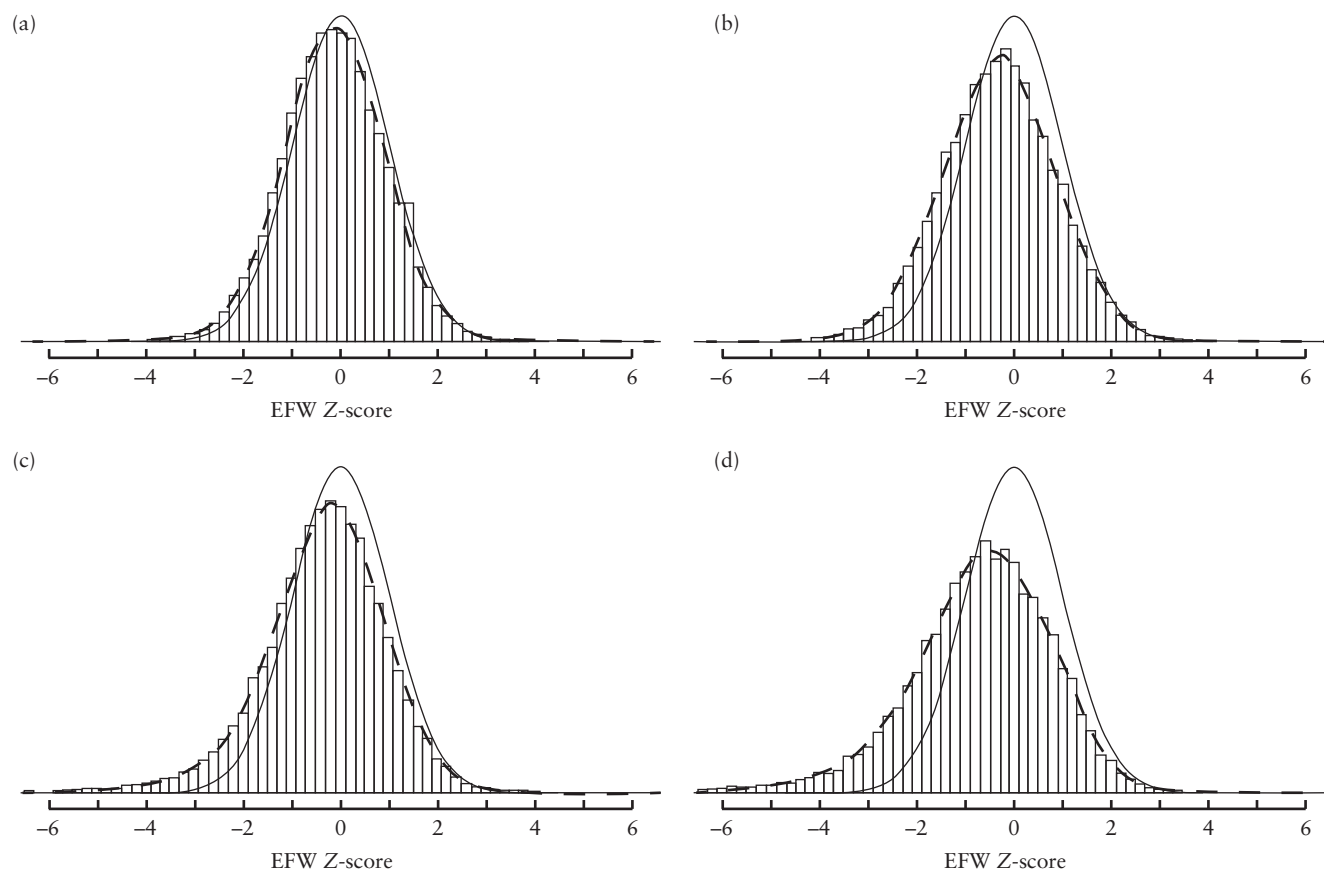


Figure 3 Distributional properties of Fetal Medicine Foundation models in Danish validation cohort of dichorionic twin pregnancies resulting in two liveborn children: (a) scheduled visits and delivery at or after 37 + 0 weeks; (b) non-scheduled visits and delivery at or after 37 + 0 weeks; (c) scheduled visits and delivery before 37 + 0 weeks; and (d) non-scheduled visits and delivery before 37 + 0 weeks. Smooth curve represents standard Gaussian distribution with a mean of 0 and SD of 1. Dashed line depicts estimated fetal weight (EFW) Z-score distribution in validation cohort.

Table 2 Estimated fetal weight Z-scores in validation cohort, according to gestational age at delivery and type of visit

Visit	DC twin pregnancy				MCDA twin pregnancy			
	Delivery < 37 + 0 weeks		Delivery ≥ 37 + 0 weeks		Delivery < 36 + 0 weeks		Delivery ≥ 36 + 0 weeks	
	n	Z-score	n	Z-score	n	Z-score	n	Z-score
Scheduled	22 756	-0.35 ± 1.31	42 453	-0.14 ± 1.05	10 118	-0.25 ± 1.29	18 816	-0.09 ± 1.01
Non-scheduled	11 513	-0.69 ± 1.50	18 624	-0.31 ± 1.17	1338	-0.63 ± 1.68	1360	-0.44 ± 1.22

Z-scores are given as mean ± SD. DC, dichorionic; MCDA, monochorionic diamniotic.

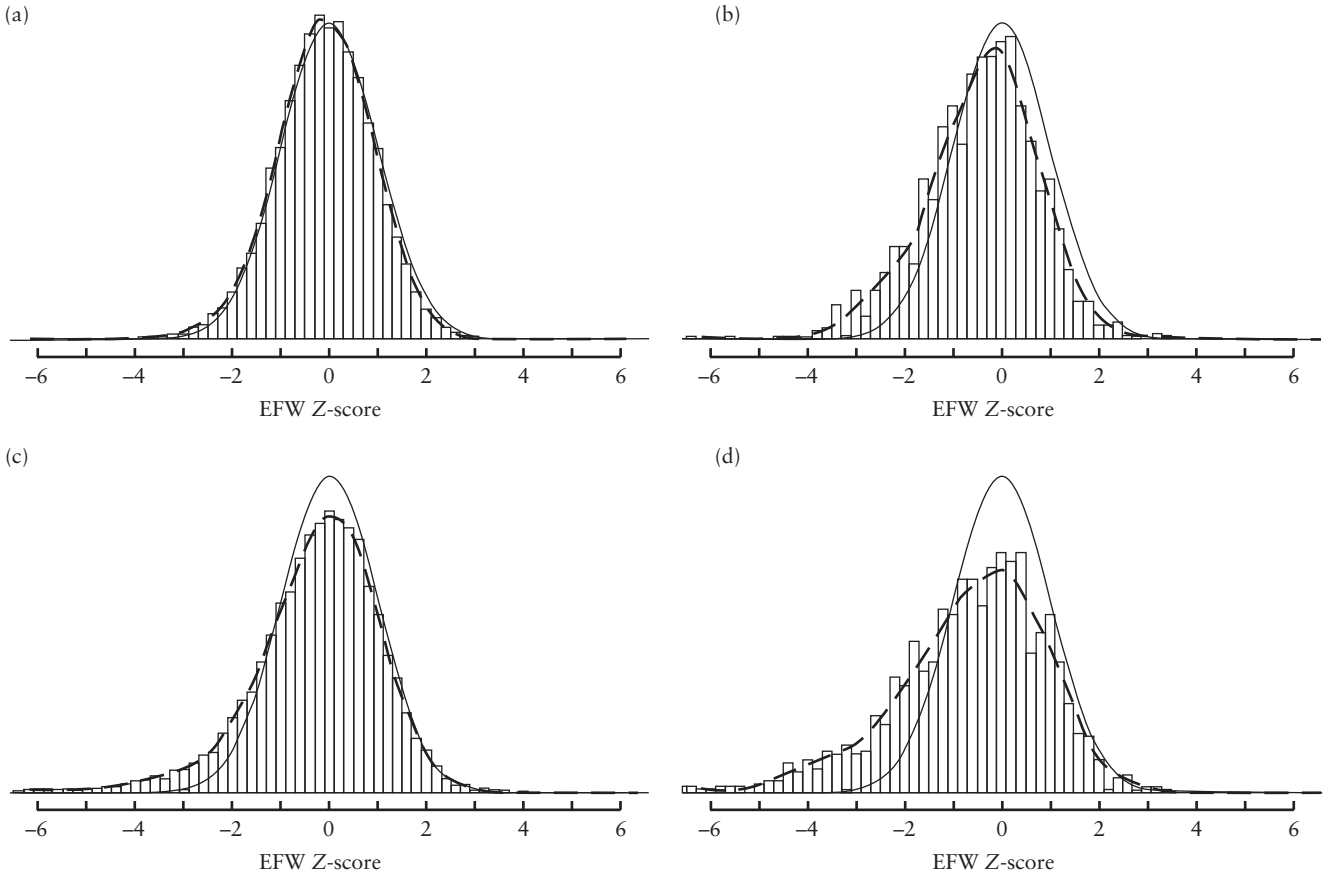


Figure 4 Distributional properties of Fetal Medicine Foundation models in Danish validation cohort of monochorionic diamniotic twin pregnancies resulting in two liveborn children: (a) scheduled visits and delivery at or after 36 + 0 weeks; (b) non-scheduled visits and delivery at or after 36 + 0 weeks; (c) scheduled visits and delivery before 36 + 0 weeks; and (d) non-scheduled visits and delivery before 36 + 0 weeks. Smooth curve represents standard Gaussian distribution with a mean of 0 and SD of 1. Dashed line depicts estimated fetal weight (EFW) Z-score distribution in validation cohort.

36 + 0 weeks or later, a larger proportion of Z-scores was observed in the lower range, indicating a left shift in the distribution (mean ± SD EFW Z-score, -0.44 ± 1.22). For pregnancies delivered before 36 + 0 weeks, weight estimates (Z-scores) for scheduled visits showed a less pronounced decrease (mean ± SD EFW Z-score, -0.25 ± 1.29), while non-scheduled visits of pregnancies delivered before 36 + 0 weeks displayed the highest proportion of small weight estimates (mean ± SD EFW Z-score, -0.63 ± 1.68). These findings were consistent with our definition of complicated MCDA pregnancy, namely one characterized by non-scheduled visits and delivery before 36 + 0 weeks. Figure S2 shows the distributional properties of Z-scores at all scheduled visits and

at each 2-weekly scheduled visit in pregnancies delivered at 36 + 0 weeks or later.

DISCUSSION

Principal findings

This study demonstrates the validity of the FMF models for chorionicity-specific fetal growth in twin pregnancies. For DC twins, the model showed favorable agreement with the validation cohort for scheduled visits in pregnancies delivered at 37 + 0 weeks or later. However, non-scheduled visits and deliveries before 37 + 0 weeks exhibited a higher proportion of smaller weight estimates.

Similar trends were observed for MCDA twins, with favorable agreement between the model and the validation cohort for scheduled visits in pregnancies delivered at 36 + 0 weeks or later. Non-scheduled visits and deliveries before 36 + 0 weeks had a greater proportion of small weight estimates.

Comparison with other studies

In the development study, Wright *et al.*¹⁰ established the FMF chorionicity-specific models for fetal growth in twin pregnancies. These models were derived from an extensive cohort of twin pregnancies pooled from four institutions across three countries. The criteria for model development included uncomplicated pregnancies and scheduled visits, as per the definitions in the present study. Following the development of the models, Wright *et al.* published a study utilizing data from 1194 twin pregnancies and 29 035 singleton pregnancies in the EVENTS trial²⁸ and investigated the use of singleton *vs* twin-specific growth charts at 36 weeks²⁹. The findings indicated that twins classified as SGA using singleton charts exhibited higher neonatal morbidity. This demonstrates the superiority of singleton charts in classifying and detecting growth-related risks, supporting their use for consistent fetal-growth classification in both twins and singletons. Other investigators recommend enhanced precision in detecting adverse outcome through the application of twin-specific growth charts^{30,31}. These studies are limited by their relatively small sample sizes, resulting in a lack of statistical power to address conclusively the relatively infrequent outcomes observed. Moreover, they fail to address the potential oversight in identifying adverse outcomes among pregnancies categorized as AGA based on the twin-specific chart but as SGA based on the singleton chart.

Previous research on fetal growth in twin pregnancies has demonstrated a lack of consistency, utilizing various methods and differing definitions for categorizing pregnancies as complicated or uncomplicated^{13,32–42}. This includes varying approaches to excluding complicated pregnancies, thus adding another layer of uncertainty and complicating direct comparisons between studies. Most importantly, to our knowledge, none of these studies has been validated externally, limiting their generalizability^{13,32–42}. There has been no prospective study to date that examined the consequences of implementing growth charts specifically designed for twins.

Clinical implications

This external validation of the FMF models for chorionicity-specific fetal growth in twins has several clinical implications. First, our study shows a strong level of agreement between the FMF models and the largest detailed national dataset of DC and MCDA twin pregnancies documented to date. This evidence supports the use of the FMF models as a novel, standardized approach for evaluating fetal growth in twin pregnancies.

This approach involves combining and interpreting chorionicity-specific fetal growth in twins relative to fetal growth in singletons. Adopting this method could potentially enhance the accuracy and scientific integrity of fetal-growth evaluation by avoiding the normalization of smallness in twins while still incorporating specific details of twin growth. Second, using the singleton percentile scale, the validated models offer an improved standardized approach for evaluating fetal growth in twins relative to singletons. This technique facilitates improved comparisons and more accurate identification of potential deviations from expected growth patterns, especially at earlier gestations. By doing so, the approach aims to reduce unnecessary visits and interventions in fetal monitoring while ensuring that cases at risk are not overlooked. However, these associations have not yet been proven and should be a focus for future research to confirm this method's effectiveness and safety. Third, the validated models improve our understanding of fetal growth in twin pregnancies by potentially aiding in the identification of abnormal growth trajectories. These trajectories do not correspond to either a twin-specific or a singleton reference but contain features of both. Identifying such trajectories could be crucial in signaling underlying complications or increased risk of adverse outcome.

Strengths and limitations

This study has several strengths. First, the external validation was performed in a national Danish cohort of, to our knowledge, the largest highly detailed dataset of DC and MCDA twin pregnancies. The study benefits from a reduced risk of selection bias, owing to the detailed and comprehensive data provided by the Danish registries. The multicenter development cohort and national validation cohort enhance the generalizability of our findings. Additionally, prenatal assessments in the validation cohort were performed in a setting with a very high participation rate covered by the Danish tax-funded public healthcare system for all pregnant women, ensuring equal access to care. The models used in this study are based on the MCMC method, effectively incorporating the relationships between interpregnancy, interfetal and intervisit factors. The reference cohort was well defined, comprising uncomplicated pregnancies resulting in two liveborn children, seen for scheduled visits and delivered at or after the internationally recommended gestational age. The study also successfully demonstrated differences in pregnancies delivered earlier than the reference cohort and those seen for non-scheduled visits.

This study is subject to certain limitations. First, it utilizes retrospective development and validation cohorts, which could introduce inherent biases. Second, the models are exclusively based on EFW calculated using the Hadlock-3 formula. Apart from their use in the Hadlock-3 formula, detailed individual biometric measurements (head circumference, abdominal circumference and femur length) for each twin are not utilized in the models, nor do the models include Doppler measurements or amniotic

fluid assessment. Third, the validation cohort consisted predominantly of Caucasian women with normal body mass index, which may limit the applicability of the findings to diverse populations. Fourth, the models used in this study, similar to the FMF models for singletons, do not account for fetal sex, a factor known to influence growth patterns. This exclusion may limit the precision of fetal-growth assessment, as male and female fetuses typically follow different growth trajectories. Fifth, chorionicity is crucial for accurate monitoring protocols but was not routinely confirmed by postpartum examination of the placenta in our cohort. In instances of unresolved uncertainty regarding chorionicity, pregnancies were managed as monochorionic, adhering to the more intensive monitoring protocol applicable to these pregnancies. Sixth, categorizing non-scheduled visits as indicative of complicated pregnancy may be an oversimplification, as these visits could result from suspected issues that warranted additional monitoring rather than confirmed complications. Last, the inclusion of different healthcare systems in the development and validation cohorts might have introduced variability in care and outcomes, potentially impacting on the generalizability of the study findings in various healthcare settings.

Conclusions

The results of this external validation study, based on a large Danish nationwide cohort of twin pregnancies, reinforce the validity of the chorionicity-specific FMF models for describing fetal growth in twin pregnancies. By juxtaposing singleton and chorionicity-specific twin growth distributions, the models provide a more nuanced and detailed evaluation of fetal growth in twins. This study provides robust evidence validating the FMF models as reliable references for understanding fetal growth in twin pregnancies.

It is imperative to establish associations between the FMF models, pregnancy outcomes and long-term outcomes for the children. Furthermore, future studies should explore the integration of Doppler measurements, biomarkers and other relevant factors of importance to fetal health and development. Through these efforts, we hope to develop more precise, evidence-based management strategies, providing the best possible health trajectory for these high-risk pregnancies and enhancing the standard of obstetric care.

ACKNOWLEDGMENT

S.E.K. is supported by the BRIDGE–Translational Excellence Programme (bridge.ku.dk) at the Faculty of Health and Medical Sciences, University of Copenhagen, funded by the Novo Nordisk Foundation (Grant agreement no. NNF23SA0087869).

REFERENCES

- Monden C, Pison G, Smits J. Twin Peaks: more twinning in humans than ever before. *Hum Reprod*. 2021;36(6):1666–1673.
- Bassil K, Shah P, Barrington K, et al. The changing epidemiology of preterm twins and triplets admitted to neonatal intensive care units in Canada, 2003 to 2008. *Am J Perinatol*. 2012;29:237–244.
- Janvier A, Spelke B, Barrington KJ. The epidemic of multiple gestations and neonatal intensive care unit use: the cost of irresponsibility. *J Pediatr*. 2011;159:409–413.
- Elliott JP, Istwan NB, Collins A, Rhea D, Stanziano G. Indicated and non-indicated preterm delivery in twin gestations: impact on neonatal outcome and cost. *J Perinatol*. 2005;25:4–7.
- Peter C, Wenzlaff P, Kruempelmann J, Alzen G, Buelmann E, Gruessner SE. Perinatal morbidity and early neonatal mortality in twin pregnancies. *Open J Obstet Gynecol*. 2013;3:78–89.
- Nicolaides KH, Wright D, Syngelaki A, Wright A, Akolekar R. Fetal Medicine Foundation fetal and neonatal population weight charts. *Ultrasound Obstet Gynecol*. 2018;52:44–51.
- Maršál K, Persson P-H, Larsen T, Lilja H, Selbing A, Sultan B. Intrauterine growth curves based on ultrasonically estimated foetal weights. *Acta Paediatr*. 1996;85:843–848.
- Papageorgiou AT, Ohuma EO, Altman DG, et al. International standards for fetal growth based on serial ultrasound measurements: the Fetal Growth Longitudinal Study of the INTERGROWTH-21st Project. *Lancet*. 2014;384:869–879.
- Morin L, Lim K. No. 260-ultrasound in twin pregnancies. *J Obstet Gynaecol Can*. 2017;39:e398–e411.
- Wright A, Wright D, Chaveeva P, et al. Fetal Medicine Foundation charts for fetal growth in twins. *Ultrasound Obstet Gynecol*. 2024;63:181–188.
- Hirsch L, Barrett J, Fox NS, Rebarber A, Kingdom J, Melamed N. Should twin-specific growth charts be used to assess fetal growth in twin pregnancies? *Am J Obstet Gynecol*. 2022;227:10–28.
- Briffa C, Stirrup O, Huddy C, et al. Twin chorionicity-specific population birth-weight charts adjusted for estimated fetal weight. *Ultrasound Obstet Gynecol*. 2021;58:439–449.
- Stirrup OT, Khalil A, D'Antonio F, Thilaganathan B, Southwest Thames Obstetric Research Collaborative (STORK). Fetal growth reference ranges in twin pregnancy: analysis of the Southwest Thames Obstetric Research Collaborative (STORK) multiple pregnancy cohort. *Ultrasound Obstet Gynecol*. 2015;45:301–307.
- Khalil A, D'Antonio F, Dias T, Cooper D, Thilaganathan B, Southwest Thames Obstetric Research Collaborative (STORK). Ultrasound estimation of birth weight in twin pregnancy: comparison of biometry algorithms in the STORK multiple pregnancy cohort. *Ultrasound Obstet Gynecol*. 2014;44:210–220.
- Astraia, GmbH, Munich, Germany. <https://www.nexus-astraia.com/>
- Ekelund CK, Petersen OB, Jørgensen FS, et al. The Danish Fetal Medicine Database: establishment, organization and quality assessment of the first trimester screening program for trisomy 21 in Denmark 2008–2012. *Acta Obstet Gynecol Scand*. 2015;94:577–583.
- Schmidt M, Pedersen L, Sørensen HT. The Danish Civil Registration System as a tool in epidemiology. *Eur J Epidemiol*. 2014;29:541–549.
- Ekelund CK, Kopp TL, Tabor A, Petersen OB. The Danish Fetal Medicine database. *Clin Epidemiol*. 2016;8:479–483.
- Stenhouse E, Hardwick C, Maharaj S, Webb J, Kelly T, Mackenzie FM. Chorionicity determination in twin pregnancies: how accurate are we? *Ultrasound Obstet Gynecol*. 2002;19:350–352.
- Sepulveda W, Sebire NJ, Hughes K, Odibo A, Nicolaides KH. The lambda sign at 10–14 weeks of gestation as a predictor of chorionicity in twin pregnancies. *Ultrasound Obstet Gynecol*. 1996;7:421–423.
- Hadlock FP, Harrist RB, Martinez-Poyer J. In utero analysis of fetal growth: a sonographic weight standard. *Radiology*. 1991;181:129–133.
- Milner J, Arezina J. The accuracy of ultrasound estimation of fetal weight in comparison to birth weight: a systematic review. *Ultrasound*. 2018;26:32–41.
- Tsakiridis I, Giouleka S, Mamopoulos A, Athanasiadis A, Dagklis T. Management of twin pregnancies: a comparative review of national and international guidelines. *Obstet Gynecol Surv*. 2020;75:419–430.
- Nicholas L, Fischbein R, Ernst-Milner S, Wani R. Review of international clinical guidelines related to prenatal screening during monochorionic pregnancies. *J Clin Med*. 2021;10:1128.
- Hamra G, MacLehose R, Richardson D. Markov chain Monte Carlo: an introduction for epidemiologists. *Int J Epidemiol*. 2013;42:627–634.
- Team R. RStudio: Integrated Development for R. <http://www.rstudio.com/>
- Kristensen SE, Ekelund CK, Sandager P, et al. Risks and pregnancy outcome after fetal reduction in dichorionic twin pregnancies: a Danish national retrospective cohort study. *Am J Obstet Gynecol*. 2023;228:590.e1–590.e12.
- Rehal A, Benkő Z, De Paco Matallana C, et al. Early vaginal progesterone versus placebo in twin pregnancies for the prevention of spontaneous preterm birth: a randomized, double-blind trial. *Am J Obstet Gynecol*. 2021;224:86.e81–86.e19.
- Wright D, Wright A, Rehal A, et al. Incidence of neonatal morbidity in small-for-gestational-age twins based on singleton and twin charts. *Ultrasound Obstet Gynecol*. 2024;63:365–370.
- Kalafat E, Sebghati M, Thilaganathan B, Khalil A, Southwest Thames Obstetric Research Collaborative (STORK). Predictive accuracy of Southwest Thames Obstetric Research Collaborative (STORK) chorionicity-specific twin growth charts for stillbirth: a validation study. *Ultrasound Obstet Gynecol*. 2019;53:193–199.
- Briffa C, Di Fabrizio C, Kalafat E, et al. Adverse neonatal outcome in twin pregnancy complicated by small-for-gestational age: twin vs singleton reference charts. *Ultrasound Obstet Gynecol*. 2022;59:377–384.
- Araujo Júnior E, Ruano R, Javadian P, et al. Reference charts for fetal biometric parameters in twin pregnancies according to chorionicity. *Prenat Diagn*. 2014;34:382–388.

33. Shivkumar S, Himes KP, Hutcheon JA, Platt RW. An ultrasound-based fetal weight reference for twins. *Am J Obstet Gynecol*. 2015;213:224.e1–9.
34. Sekiguchi M, Mikami M, Nakagawa C, et al. An ultrasonographic estimated fetal weight reference for Japanese twin pregnancies. *J Med Ultrason* (2001). 2019;46:209–215.
35. Wilkof Segev R, Gelman M, Maor-Sagie E, Shrim A, Hallak M, Gabbay-Benziv R. New reference values for biometrical measurements and sonographic estimated fetal weight in twin gestations and comparison to previous normograms. *J Perinat Med*. 2019;47:757–764.
36. Min SJ, Luke B, Gillespie B, et al. Birth weight references for twins. *Am J Obstet Gynecol*. 2000;182:1250–1257.
37. Stirrup OT, Khalil A, D'Antonio F, Thilaganathan B. Patterns of second- and third-trimester growth and discordance in twin pregnancy: analysis of the Southwest Thames Obstetric Research Collaborative (STORK) multiple pregnancy cohort. *Fetal Diagn Ther*. 2017;41:100–107.
38. Gabbay-Benziv R, Crimmins S, Contag SA. Reference values for sonographically estimated fetal weight in twin gestations stratified by chorionicity: a single center study. *J Ultrasound Med*. 2017;36:793–798.
39. Ong S, Lim M-N, Fitzmaurice A, Campbell D, Smith AP, Smith N. The creation of twin centile curves for size. *BJOG*. 2002;109:753–758.
40. Grantz KL, Grewal J, Albert PS, et al. Dichorionic twin trajectories: the NICHD Fetal Growth Studies. *Am J Obstet Gynecol*. 2016;215:221.e1–221.e16.
41. Ghi T, Prefumo F, Fichera A, et al. Development of customized fetal growth charts in twins. *Am J Obstet Gynecol*. 2017;216:514.e1–514.e17.
42. Hirsch L, Okby R, Freeman H, et al. Differences in fetal growth patterns between twins and singletons. *J Mat-Fet Neonat Med*. 2020;33:2546–2555.

SUPPORTING INFORMATION ON THE INTERNET

The following supporting information may be found in the online version of this article:



Figure S1 Distributional properties of Fetal Medicine Foundation models in Danish validation cohort of dichorionic twin pregnancies delivering two liveborn children at or after 37 + 0 weeks' gestation, at each of the 4-weekly scheduled visits. Smooth curve represents standard Gaussian distribution with mean of 0 and SD of 1. Dashed red line depicts Z-score distribution in validation cohort.

Figure S2 Distributional properties of Fetal Medicine Foundation models in Danish validation cohort of monochorionic diamniotic twin pregnancies delivering two liveborn children at or after 36 + 0 weeks' gestation, at each of the 2-weekly scheduled visits. Smooth curve represents standard Gaussian distribution with mean of 0 and SD of 1. Dashed red line depicts Z-score distribution in validation cohort.