

Developmental instability and its relationship to mental health in two historic Dutch populations

Hagg, Alieske Christiene^a; Van der Merwe, Alie Emily^b, Steyn, Maryna^c

^aDepartment of Anatomy, Faculty of Health Sciences, University of Pretoria, 09 Bophelo Road, Pretoria 0001, South Africa, HaggA@saps.gov.za

^bDepartment of Anatomy, Embryology and Physiology, Academic Medical Centre, Meibergdreef 15, 1100 DD Amsterdam, The Netherlands, a.e.vandermerwe@amc.uva.nl

^cHuman Variation and Identification Research Unit, School of Anatomical Sciences, Faculty of Health Sciences, University of the Witwatersrand, 7 York Road, Parktown, 2193, Johannesburg, South Africa, maryna.steyn@wits.ac.za

Corresponding author: Hagg, Alieske Christiene

ABSTRACT

This study aimed to assess the magnitude and patterns of fluctuating asymmetry as an indicator of developmental instability between two urban archaeological Dutch populations. The sample comprised of 209 adult individuals representing the general population of Alkmaar, the Netherlands, dating to the 18th to early 19th century (Grote Kerk, n=134), and a psychiatric hospital sample of the 19th to early 20th century (Meerenberg, n=75). Fluctuating asymmetry was assessed from left and right measurements recorded from various traits on the cranium, mandible, and dentition. Three non-specific skeletal indicators of stress were documented to aid in the interpretation of the differences in asymmetry. No significant difference in developmental instability, as reflected by fluctuating asymmetry, was apparent between the two populations. However, individuals who presented with skeletal lesions indicative of stress were significantly more asymmetric than individuals who did not present with any of the lesions. The observed frequencies of the pathological changes and socio-economic history suggest that the two populations experienced similar levels of stress, even though the source and duration of the stress might have been different. The possibility that the mentally institutionalized are not as developmentally unstable as suggested by previous research should be considered.

Key words

Fluctuating asymmetry, environmental stress, genetic stress, canalization

1. INTRODUCTION

Developmental homeostasis and two of its components, developmental stability and canalization, are important for the attainment of bilaterally symmetrical anatomical structures during ontogeny (Møller and Swaddle, 1997). Developmental stability refers to the inherent ability of an organism to resist random errors or stressors under specific internal (genetic) and external (environmental) conditions (Møller and Swaddle, 1997; Palmer, 1994). Bilateral structures or traits, such as the left and right orbits, are products of a common gene complex with identical developmental pathways. Should developmental homeostasis be maintained throughout development, the phenotypic expression of this identical genotype will result in identical phenotypes that are bilaterally symmetrical (Palmer and Strobeck, 2003; Van Valen, 1962). However, when changes in the internal and external conditions of development occur, canalization acts as a buffer or regulator to these changes to ensure that developmental homeostasis is maintained and that the intended developmental pathway is followed (Møller and Swaddle, 1997; Waddington, 1957, 1942). The inability to maintain homeostasis due to developmental instability or ineffective canalization will lead to bilateral asymmetry of a structure or trait (Van Valen, 1962).

Three types of bilateral asymmetry are described in the literature: fluctuating asymmetry (FA)¹, directional asymmetry (DA) and antisymmetry. Antisymmetry and DA of a trait are usually associated with behavioral factors, also referred to as adaptive asymmetries. In DA a morphological trait in a population will express a constant greater increase in development on either the right or the left side of the plane of symmetry, whereas in antisymmetry, the preference in increased size towards any one side within a population will be random, resulting in bimodal symmetry around a mean of zero (Palmer, 1994; Van Valen, 1962). Fluctuating asymmetry, on the other hand, has mainly been linked to developmental instability, and is defined as the morphological inequality of bilateral anatomical structures such that the average expression of individual asymmetries per trait in a population will be symmetrical (Møller and Swaddle, 1997; Palmer, 1994; Van Valen, 1962). This paper will focus on FA and its utilization as an indicator of developmental instability within past populations.

While FA has been utilized in a vast number of studies as an indicator of developmental instability, its reliability for such purposes has been questioned (Bjorksten et

¹Abbreviations: FA: Fluctuating asymmetry; DA: Directional asymmetry; GRK: Grote Kerk population; MeB: Meerenberg population; ME: Measurement error

al., 2000; e.g. Black, 1980; Gawlikowska et al., 2007). The support for its use as an indicator of developmental instability seems to be based on positive relationships between FA and environmental, genetic, and other indicators of stress. The majority of literature on FA in human archeological remains, specifically on the cranium and dentition, revealed increased levels of FA that coincided with lower socio-economic status, malnutrition, higher mortality rates, retarded growth during late childhood, and higher frequencies of enamel hypoplasia (EH) and cribra orbitalia (Bailit et al., 1970; DeLeon, 2007; Guatelli-Steinberg et al., 2006; Harris and Nweeia, 1980; Perzigian, 1977; Storm, 2007; Weisensee, 2013). For example, Storm (2007, 2008) found increased levels of FA in human archaeological remains from the Medieval to the Victorian periods in England that coincided with the environmental and social changes brought on by the British Industrial Revolution. Additional comparative analyses, based on the socio-economic status (SES) of a number of the archaeological populations, revealed higher levels of FA in the populations of lower SES, further supporting the premise that FA, and inherently developmental instability, increases with adverse external or environmental conditions. However, in other studies, the relationship was either not significant (Arnqvist and Thornhill, 1998; Bjorksten et al., 2000; Gawlikowska et al., 2007), or only evident for a minority of traits (Hoover and Matsumura, 2008).

The positive relationship between FA and levels of stress is based on the premise that various external and internal factors, termed developmental noise or stress, can negatively and additively affect developmental homeostasis (Møller and Swaddle, 1997; Van Valen, 1962). Complete bilateral symmetry will only occur when developmental homeostasis sufficiently buffers the developmental stresses during development, or in other words when the developmental pathways are highly canalized. When the developmental stress or stressors are greater than the threshold level or buffering ability of a trait, the original phenotypic expression of the genotype will not be reached, resulting in bilateral asymmetry or FA of a trait. It follows that the less canalized a certain developmental pathway or the greater the developmental instability, the greater the amount of FA will be (Adams and Niswander, 1967; Palmer and Strobeck, 2003; Storm, 2008; Waddington, 1942).

Chronic periods of certain environmental stressors, such as malnutrition or nutritional deficiencies, infection, and other factors or illnesses related to low SES or economic hardship, are known to cause pathological lesions on dry bone and dentition, such as periostitis, EH, cribra orbitalia and porotic hyperostosis (Mann and Murphy, 1990; Reitsema and McIlvaine, 2014). However, some of these indicators of stress, such as EH, have also been associated with syndromes and congenital disorders, and can be attributed to the

mutation of several developmental regulatory genes, or in other words, internal or genetic stress (Brook, 2009; Schuurs, 2013; Seow, 1997).

A factor that has been linked to decreased developmental stability in a group of individuals relative to the rest of the population is mental and neurological deficiencies or disorders. A vast number of studies (Malina and Buschang, 1984; Markow and Gottesman, 1989; Markow and Wandler, 1986; Martin et al., 1999; Reilly et al., 2001; Shackelford and Larsen, 1997) have associated various mental disorders and neurological impairments in living individuals with an increase in developmental instability and increased levels of asymmetry. Cerebral palsy, schizophrenia, and depression in men are some of the disorders that have been associated with decreased developmental stability. Perhaps the most relevant asymmetry study on the relationship between mental status and FA was conducted by Malina and Buschang (1984) on living males. In this study, eleven bilateral measurements were collected on 200 males with a history of mental disorders and/or neurological impairments (aged nine to 52 years), and on 202 mentally healthy males ranging from five to 35 years of age. The comparison revealed an overall higher magnitude of asymmetry within the males with mental disorders, with significant differences in the upper limb measurements.

Due to their alleged decreased developmental stability, it can be speculated that individuals with a history of mental disorders or neurological impairments will exhibit greater magnitudes of FA relative to the general ('mentally healthy') population when subjected to equal types and levels of stress. A search of the literature revealed no previous studies between FA and mental health in skeletal samples.

In light of the abovementioned hypothesis that individuals suffering from a mental or neurological disorder will exhibit greater levels of FA relative to the general population, this study aimed to compare the level of developmental stability and stress between two Dutch archaeological populations by means of an assessment of the magnitude of FA in the cranium, mandible and dentition. One population represents the general population, while the other population represents individuals buried in the graveyard of a former Dutch mental institution.

2. MATERIALS AND METHODS

2.1 Skeletal samples

This study was conducted with research and ethical approval obtained from the Academic Medical Centre (AMC), University of Amsterdam, and the Faculty of Health Sciences Research Ethics Committee, University of Pretoria (393/2014), respectively. The skeletal sample investigated comprised of 209 adult individuals from two Dutch

archaeological populations, 134 from Grote Kerk (GRK) and 75 from Meerenberg (MeB). Both skeletal collections are housed at the AMC in Amsterdam, the Netherlands.

The GRK sample dates to the 18th and early 19th century and was excavated from the church floor of the St. Laurenskerk in Alkmaar, the Netherlands. Burial underneath church floors was customary in the Netherlands from the Middle Ages to the late 1820's, after which it was banned by law. Church burials were substantially more expensive than burials within graveyards and were mainly a custom that families of middle to high SES could afford. However, burial within the church of Alkmaar only became excessively expensive by the end of the 18th century (Baetsen, 2001; Baetsen et al., 1997; Bitter, 2002).

The MeB sample was excavated from a Catholic cemetery of the historic Meerenberg psychiatric hospital in Bloemendaal, the Netherlands. Sixty-six graves were exhumed during a rescue excavation after the accidental disturbance of the graves by property development in 2009. The graves dated from 1891 to 1914, and the individuals within the sample likely lived during the early 19th to the early 20th century (Van der Merwe et al., 2013). The Meerenberg psychiatric hospital was established during the 1840's. During the 19th century, Meerenberg was one of the largest and most prestigious mental asylums in the Netherlands, housing chronic patients with more advanced levels of mental disorders or neurological impairments relative to the other mental asylums (Goldberg and Graham, 2013; Van Twuyver, 2000; Vijselaar, 1982). Both patients and staff members were buried in the cemeteries on the Meerenberg grounds (Van Twuyver, 2000).

2.2 Methods

Twenty-four bilateral measurements from the cranium and mandible, and two bilateral measurements on all permanent teeth, except the third molars, were chosen for an evaluation of the developmental instability between the two populations. The measurements and their definitions are indicated in Figure 1.

All measurements were based on previously defined measurements and procedures for dry bone (Buikstra and Ubelaker, 1994; Howells, 1973; Storm, 2009) and dentition (Guatelli-Steinberg et al., 2006; Kieser, 1990), except for three cranial measurements specifically defined for FA analyses by Storm (2009). The latter measurements include mastoid length (MPL), mastoid breadth (MPB) and mandibular length (MAL). The measurements were recorded with the use of sliding- (to the nearest 0.01 mm) and spreading (to the nearest 0.1 mm) calipers. Not all measurements were recorded on every individual due to the incomplete or fragmentary condition of the remains. Measurements recorded on elements exhibiting taphonomic, traumatic or pathologic changes were only included if it did

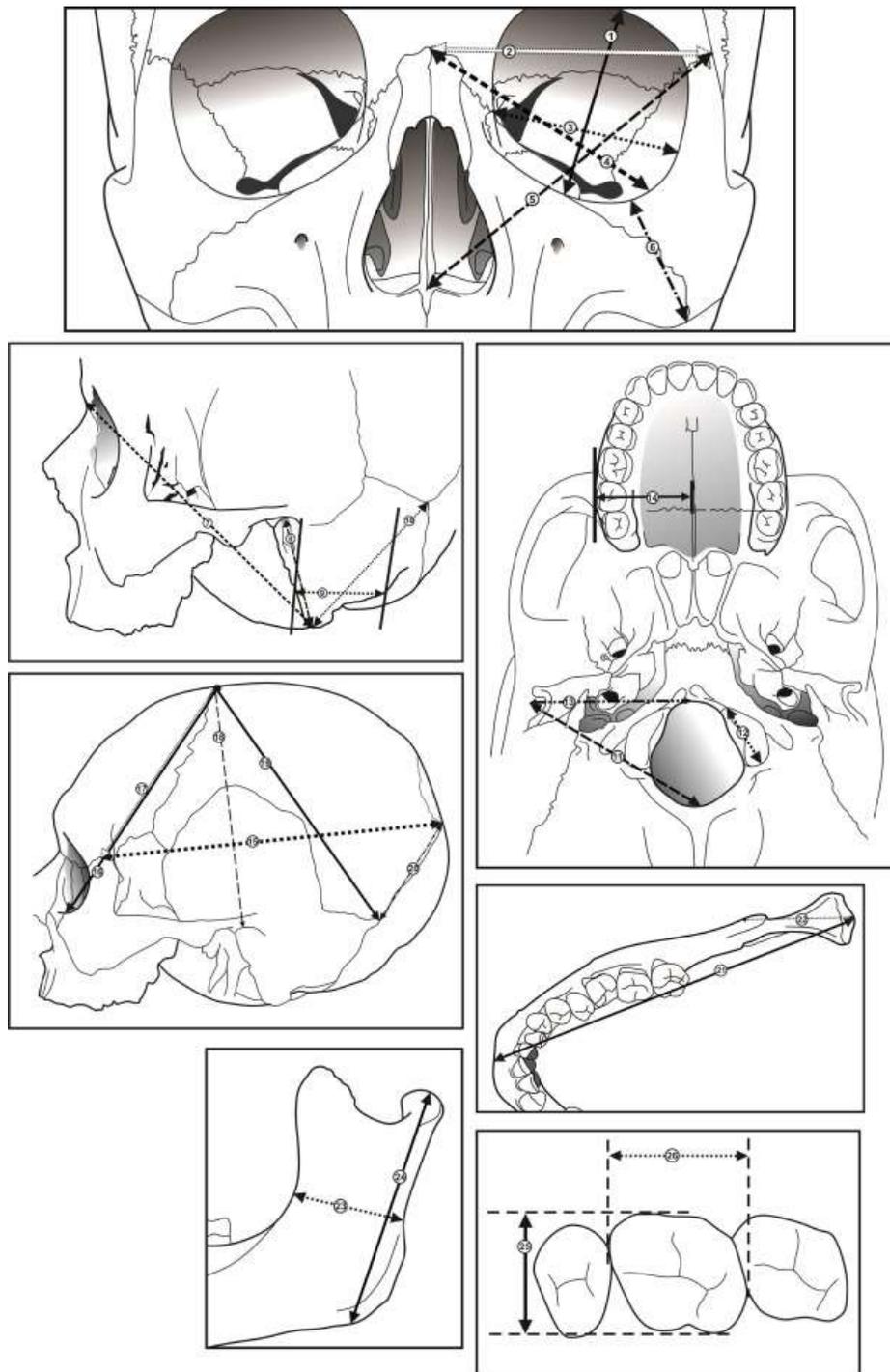


Figure 1. Abbreviations and description of recorded measurements (Adapted from Kieser, 1990; and Storm, 2009): 1. COBH: Orbital height; 2. CFMTN: Chord frontomalare-nasion; 3. COBB: Orbital breadth; 4. CNOR: Diagonal orbital breadth; 5. CFMTNS: Chord frontomolare-nasospinale; 6. CMAH: Malar height; 7. CNMS: Chord nasion-mastoidale; 8. CMPL: Length of mastoid process; 9. CMPB: Breadth of mastoid process; 10. CMSAST: Chord mastoidale-asterion; 11. COPO: Chord opisthion-porion; 12. COCL: Length of occipital condyle; 13. CBAPO: Chord basion-porion; 14. CECMIS: Chord ectomalare-intermaxillary suture; 15. CLFMT: Chord lambda-frontomalare; 16. CBZO: Chord bregma-zygoorbitale; 17. CFMTB: Chord frontomalare-bregma; 18. CBPO: Chord bregma-porion; 19. CBAST: Chord bregma-asterion; 20. CLAST: Chord lambda-asterion; 21. MAL: Mandibular length; 22. MXRB: Maximum ramus breadth; 23. MRH: Maximum ramus height; 24. MIRB: minimum ramus breadth; 25. BL: Maximum buccolingual diameter; 26. MD: Maximum mesiodistal diameter.

not affect the measurement. Teeth with excessive attrition, carious lesions, restorations, fractures or marked calculus deposits were not measured if it affected the maximum buccolingual (BL) or mesiodistal (MD) diameters of a tooth (Guatelli-Steinberg et al., 2006; Kieser, 1990). Nevertheless, the majority of measurements could be recorded on all individuals.

2.2.1 Demographic profile

Adult sex was mainly estimated by means of the assessment of pelvic (Klaes et al., 2012; Phenice, 1969) and cranial (Buikstra and Ubelaker, 1994; Walker, 2008) morphological traits, as well as metrically by means of linear discriminant analyses (LDA) of the long bone dimensions (Morrison, 1969). Linear discriminant analyses were utilized where sex estimation was ambiguous or unclear using only the morphological techniques (due to incomplete remains or the lack of the expression of distinctive sexual characteristics). As no reliable population specific data exist, the LDA was conducted against the measurements of individuals within the population who could easily be sexed by means of the morphological techniques. Individuals for whom sex could not be estimated remained classified as 'indeterminate' sex.

Age-at-death was assessed through age related changes in the pubic symphyseal surfaces (Brooks and Suchey, 1990; Hartnett, 2010), sternal rib ends (İşcan et al., 1985, 1984a, 1984b), and cranial suture closure (Acsádi and Nemeskéri, 1970; Meindl and Lovejoy, 1985). The estimation of age-at-death in this study was used only for the purpose of placement into broad age categories. Following suggestions by Storm (2009) and Falys and Lewis (2011), in order to allow for comparisons with previous and future studies, adults were divided into one of three age sub-categories: young adult (YA; 20-34 years of age), middle adult (MDA; 35-45 years of age) or mature adult (MA; > 46 years of age). Where it was not possible to assign a specific age range to an adult individual (due to incomplete or fragmentary remains), the individual was categorized simply as adult (A; > 20 year of age). The individuals were categorized into the abovementioned categories by means of average age as calculated from the estimated age ranges.

Age and sex were estimated for all individuals. Sex and age data of the GRK population were also correlated against available information from burial registers of the church. All individuals were deemed to be of Dutch origin.

2.2.2 Pathological lesions

Three general macroscopic pathological lesions were assessed on each skeleton as additional indicators of environmental stress: cribra orbitalia, subperiosteal bone reactions

(periostitis) on the long bones and alveolar processes, and EH (Brickley and Ives, 2008; Mann and Murphy, 1990; Ortner, 2003; Waldron, 2009). All pathological lesions were assessed by a single observer.

The evaluation of the skeletal and dental markers of pathology was not the main objective of the study and was only assessed on a present-absent basis. Therefore, the severity and state (active or healed) of the lesions were not taken into consideration, and any degree of expression was noted as the lesion being present. Both linear horizontal lines or grooves and pitting of the tooth enamel were considered in the scoring of EH. The frequency and stage(s) of dental development at which an insult occurred was not assessed. It must also be noted that the dentition of several individuals within the populations was affected by what is believed to be 'pipe smoker's attrition,' which could possibly have affected the scoring of EH.

2.3 Statistical methods

Before the calculation of FA, statistical analyses were conducted to ensure that the FA values were not confounded by outliers, antisymmetry, DA or measurement error. Outliers were detected and removed by means of Grubbs' outlier test (Grubbs, 1969) and visual assessments of boxplots and histograms (Laurikkala et al., 2000), while the normality of left-right distributions was assessed by means of Kolmogorov-Smirnov tests and student's *t*-tests (Bai and Ng, 2005; Massey, 1951). The presence of platykurtosis was examined for an indication of antisymmetry (Møller and Swaddle, 1997; Palmer, 1994).

The repeatability of measurements, or measurement error (ME), was assessed for each measurement individually (left and right), as well as for the difference between the sides (signed asymmetry score; right - left). The latter estimate is necessary because the repeatability of FA does not depend on the repeatability of trait size measurements (right or left), but rather on the repeatability of the difference between the left and right sides (Fields et al., 1995). Forty-five individuals were randomly selected from the entire sample for a minimum of 10 repeated bilateral measures of each measurement. All measurements were recorded by a single observer. Repeated measures were collected on different days than the first measurements, and were conducted at relatively regular intervals throughout the recording of the measurements (generally three or more days apart).

A percentage technical error of measurement (%TEM) evaluated the repeatability of each left and right measurement for a trait (Dahlberg, 1940; Norton and Olds, 1996; Perini et al., 2005; Ulijaszek and Kerr, 1999), while a two-way side by side individual analysis of variance (ANOVA) evaluated the repeatability of the signed asymmetry scores (right - left). Because FA is, in essence, an estimation of variance, any ME can artificially inflate this variance.

Therefore, in order for the resultant asymmetry to be due to more than just ME, the between-sides difference due to FA needs to be significantly greater than the between-sides differences due to ME (Palmer, 1994; Palmer and Strobeck, 2003, 1986). The ANOVA procedure assessed whether the between-side difference due to ME is smaller than the between-side differences (right - left) due to the FA measured. Based on the two-way ANOVA, measurements for which the raw asymmetry score was not significantly higher ($p < 0.05$) relative to ME, were removed from further analyses. Furthermore, a trait for which the %TEM was greater than 5.0% was deemed unacceptable (Perini et al., 2005), and was removed from further analyses. It is important to note that there is no general level of acceptance for TEM and %TEM in the literature, as it has been shown to be, amongst others, population and measurement type dependent (Ulijaszek and Kerr, 1999). Generally, a lower %TEM indicates a better correlation between the repeated measures (Geeta et al., 2009; Perini et al., 2005; Ulijaszek and Kerr, 1999).

The presence of DA will artificially inflate the FA of a trait. The possible confounding effect of DA to the observed FA was assessed by means of two analyses. Firstly, one sample student's *t*-tests were utilized to test whether the mean DA of a trait differed significantly from zero, or in other words, if the asymmetry distribution is one-sided (consistently greater to one side; Palmer, 1994; Palmer and Strobeck, 2003). Secondly, the mean difference between the right and left measurements for each trait [$DA = \ln(R_j/L_j)$] was compared to the average deviation around the mean. A trait with an average deviation around mean (right - left) larger than the mean (right - left) could be argued to mainly exhibit developmental instability, and the effect of DA could be assumed as minimal (Palmer and Strobeck, 2003).

In order to eliminate size dependence and to allow for comparisons between traits of different sizes, the FA values for each trait (measurement) were calculated with the log transformation of the asymmetry formula $FA = |R-L|/[(R+L)/2]$, namely $FA = |\ln(R_j/L_j)|$, where R_j and L_j are the measurements taken on the right and left sides of each trait respectively (Palmer, 1994; Palmer and Strobeck, 2003). The log transformation of the standard formulae also enables the FA values to be averaged for multiple traits in an individual by means of the following formula: $FA_{index} = \Sigma |\ln(R_j/L_j)| / T$, where R_j and L_j are the measurements taken on the right and left sides of each trait respectively, and T is the total number of available traits per index (Palmer and Strobeck, 2003). Table 1 contains a list of all the FA indices that were calculated per individual.

Table 1. Fluctuating asymmetry indices for multiple traits

Indices	Included measurements
Individual	All skeletal and dental measurements
Dentition	All dental measurements
Cranium	All cranial measurements
Cranium: Orbit	COBB, COBH, CNOR
Cranium: Facial	CFMTN, CFMTNS, CMAH, CECMIS, CFMTB, CBZO
Cranium: Temporal	CMPL, CMPB, CMSAST
Cranium: Base	COCL, COPO, CBAPO, CNMS
Cranium: Vault	CBPO, CBAST, CLFMT, CLAST
Mandible	All mandibular measurements

2.3.1 Comparisons of FA

The tests for normality revealed that the majority of traits were skewed and/or non-normal. In addition, the calculated FA values are given in absolute values (truncated at zero), and thus the distribution of FA values is often skewed to the right. Therefore, non-parametric tests were utilized for all comparisons of asymmetry (Palmer and Strobeck, 1986).

Non-parametric Mann-Whitney U (Wilcoxon rank sum) tests and Kruskal-Wallis ANOVA's were utilized to test for differences between the sexes and age categories respectively, within each population and with the populations pooled. Where no sex or age differences were apparent, the sexes and age groups were pooled for further analyses. Indeterminate sex and 'A' adult category were not included in the analyses between the sexes and between the adult age categories, respectively. Differences in FA magnitude between the two populations were assessed by means of Mann-Whitney U tests for each individual trait (Figure 1) and FA index (Table 1). Chi-squared tests (Rao and Scott, 1984) assessed whether significant differences existed in the prevalence of the observed macroscopic pathological conditions in each population, with the pathological changes pooled and separately per pathological condition. The median FA value per population for only the individuals who presented with at least one of the three conditions was compared between the populations by means of a Mann-Whitney U test. Lastly, a Mann-Whitney U test was also used to assess whether significant differences existed in the median FA value of the individuals who exhibited at least one pathological condition (populations pooled), and the rest of the sample (individuals with none of the three conditions).

The level of significance for all statistical tests was chosen as $p < 0.05$ (α at 0.95), with a Holm's (Holm, 1979) adjustment for multiple comparisons of the mean for the Grubb's outlier test and all statistical analyses between groups. A Holm's adjustment adjusts

the p-values for multiple comparisons of the means by means of a Bonferroni correction. It reduces the size of α for each comparison, increasing the difference needed to be significant and decreasing the chance of a type I error (Dawson and Trapp, 2004; Holm, 1979).

While the use of a large number of independent tests is not ideal (Palmer, 1994), it is important that each trait is assessed individually, as well as combined (indices) for all the possible confounding factors and comparisons of FA. Because different traits within an individual tend to differ in their buffering ability to the type and duration of stress (Storm, 2009; Van Valen, 1962), the developmental instability of each trait needs to be considered before the traits are combined for an estimation of the developmental instability of a certain region or index (for e.g. the cranium) of an individual within a population. The combination of traits enable an increase in the confidence with which the estimation of individual FA or developmental instability is measured (Palmer, 1994; Palmer and Strobeck, 2003).

3. RESULTS

The final sample comprised of 134 adult individuals from the GRK and 75 from the MeB population. Table 2 indicates the sample sizes included in the study per population, sex and age group.

Table 2. Sample sizes included in the study per population, sex and age category

	Grote Kerk				Meerenberg			
	M	F	I	Total	M	F	I	Total
YA	0	5	1	6	1	2	0	3
MDA	48	35	1	84	10	10	0	20
MA	9	29	1	39	6	18	2	26
A	1	4	0	5	3	11	12	26
Total	58	73	3	134	20	41	14	75

M=male, F=female, I=indeterminate; YA=young adult, MDA=middle adult, MA=mature adult, A=adult

The analysis of ME revealed that all but one measurement, namely the breadth of the mastoid process (CMPB; 5.58%), fell within the acceptable range for skilled intraobserver repeatability (Perini et al., 2005), with a %TEM ranging from 0.29 to 5.58% (Table 3). The two-way ANOVA indicated two skeletal traits, breadth of the mastoid process (CMPB) and ectomalare-intermaxillary suture length (CECMIS), and 18 dental traits to exhibit significantly higher ME relative to the between side difference due to FA, as indicated by no statistical significance. However, the %TEM for the dental measurements ranged from 1.3 to 4.16%, which is within the acceptable ranges for skilled observers.

Table 3. Measurement error results from technical error of measurement (TEM) values and a two-way side by individual analysis of variance (ANOVA)

Trait	N	TEM	Mean	%TEM	Sides x Individuals ANOVA			
					df	MS _{int}	F	p-value
COBB	34	0.39	39.51	1.00	15	0.670	4.129	<0.001**
COBH	38	0.36	33.95	1.07	17	0.785	5.852	<0.001**
CNOR	38	0.48	54.18	0.88	17	0.911	3.842	<0.001**
CFMTN	40	0.32	54.87	0.59	17	0.356	3.579	<0.001**
CFMTNS	36	0.41	75.99	0.54	16	2.369	14.739	<0.001**
CMAH	40	0.23	22.16	1.05	17	1.356	27.233	<0.001**
CMPL	30	0.88	29.12	3.01	12	5.193	5.941	<0.001**
CMPB	34	1.59	27.18	5.85	14	4.573	2.028	0.051
CMSAST	28	0.43	48.44	0.88	11	4.387	23.369	<0.001**
COCL	38	0.40	25.17	1.59	16	5.235	33.630	<0.001**
COPO	36	0.51	69.10	0.73	16	6.073	22.902	<0.001**
CBAPO	38	0.53	58.51	0.91	17	3.366	11.260	<0.001**
CECMIS	28	0.40	29.04	1.37	11	0.558	3.344	0.006*
CNMS	28	0.69	123.95	0.56	12	4.281	10.355	<0.001**
CBZO	28	0.64	130.83	0.49	12	2.460	5.562	<0.001**
CFMTB	36	0.52	114.26	0.46	16	4.157	14.878	<0.001**
CBPO	35	0.77	123.83	0.63	15	4.516	7.451	<0.001**
CBAST	30	0.47	133.72	0.35	13	2.987	17.154	<0.001**
CLFMT	36	0.54	169.53	0.32	16	2.254	7.963	<0.001**
CLAST	32	0.93	89.41	1.04	14	8.406	9.556	<0.001**
MAL	34	0.34	115.34	0.29	12	6.979	64.190	<0.001**
MRH	30	0.99	56.40	1.76	10	12.003	10.140	<0.001**
MXRB	36	0.44	40.64	1.08	14	2.616	19.205	<0.001**
MIRB	42	0.12	27.99	0.43	14	1.741	94.174	<0.001**
i2l_bl	24	0.07	6.24	1.13	10	0.020	1.764	0.128
i2l_md	18	0.08	5.83	1.33	8	0.030	1.457	0.241
c1l_bl	30	0.12	7.70	1.51	11	0.022	4.355	0.001**
c1l_md	25	0.10	6.46	1.47	8	0.014	2.307	0.067
pm1l_bl	30	0.12	7.68	1.50	14	0.016	1.198	0.326
pm1l_md	34	0.11	6.61	1.73	11	0.014	1.551	0.175
pm2l_bl	30	0.16	8.09	1.96	13	0.065	5.282	<0.001**
pm2l_md	32	0.17	6.97	2.37	15	0.049	3.714	<0.001**
m1l_bl	20	0.33	10.32	3.23	14	0.043	1.663	0.122
m1l_md	22	0.30	10.75	2.79	14	0.107	3.663	<0.001**
m2l_bl	20	0.20	10.09	1.97	5	0.016	0.660	0.660
m2l_md	26	0.21	10.55	2.00	6	0.089	4.545	0.009*
i1u_bl	32	0.10	6.93	1.38	10	0.042	1.067	0.430
i1u_md	26	0.07	8.19	0.88	12	0.121	2.719	0.016*
i2u_bl	26	0.11	6.21	1.81	12	0.032	3.849	0.002**
i2u_md	18	0.07	6.59	1.11	9	0.057	10.431	<0.001**
c1u_bl	26	0.18	8.20	2.16	11	0.091	6.768	<0.001**

c1u_md	26	0.19	7.58	2.46	7	0.069	15.300	<0.001**
pm1u_bl	28	0.14	8.78	1.62	12	0.033	1.054	0.434
pm1u_md	30	0.16	6.67	2.39	12	0.050	1.434	0.213
pm2u_bl	29	0.12	9.00	1.31	13	0.016	0.748	0.703
pm2u_md	30	0.18	6.50	2.78	13	0.037	1.961	0.066
m1u_bl	20	0.35	10.64	3.32	13	0.065	4.542	<0.001**
m1u_md	22	0.44	10.54	4.16	13	0.053	1.517	0.173
m2u_bl	22	0.20	10.77	1.90	5	0.010	0.739	0.609
m2u_md	26	0.32	9.94	3.24	6	0.068	0.354	0.896
i1l_bl	30	0.11	5.61	1.90	9	0.278	8.030	<0.001**
i1l_md	26	0.15	5.31	2.83	11	0.277	2.488	0.030*

*p < 0.05; **significant after a Holm's adjustment, p < 0.05; N=number of traits, df = degrees of freedom, $MS_{int}=MS_{interaction}$, sides by individual expected mean squares, F=test statistic of two-way ANOVA, i=incisor, c=canine, pm=premolar, m=molar, l=lower, u=upper, bl=buccolingual, md=mesiodistal

Based on the presence of high measurement error for both FA values and left/right sides, the following skeletal traits were removed from further analyses: breadth of the mastoid process (CMPB), and ectomalare-intermaxillary suture length (CECMIS). This follows the practices followed in similar studies (DeLeon, 2007; Hoover et al., 2005; Storm, 2009), as was suggested by Palmer and Strobeck (2003).

Only one trait, the diagonal breadth of the orbit (CNOR), exhibited significant DA, although no traits exhibited a mean asymmetry value greater than the variation around the mean. Therefore, CNOR could not be considered to exhibit significant directional asymmetry (Palmer and Strobeck, 2003), and was retained for subsequent analyses.

When considering all sex and age categories, respectively, no statistically significant differences in the magnitude of FA were evident for any of the independent traits, nor for any of the indices, both within each population and for the populations combined. Therefore, the sex and age categories were combined for all subsequent analyses.

The GRK and MeB median asymmetry per individual was almost identical in magnitude at 2.40% ($\bar{x} = 2.5\%$, $\sigma = 0.7\%$) for GRK and 2.39% ($\bar{x} = 2.6\%$, $\sigma = 0.8\%$) for MeB (Table 4). No traits or indices differed significantly between the two populations ($p > 0.05$).

Despite no observable significant differences between the populations, the mean asymmetry values of the traits and indices revealed some general patterns. For example, compared to the other regions of the cranium, the majority of facial traits (CBZO, CFMTNS, CFMTN and CFMTB) exhibited some of the lowest mean FA values, while the majority of traits located in the cranial base (COCL, COPO, and CBAPO), exhibited some of the highest

mean levels of FA. The median difference between the two cranial indices was statistically significant ($P=0.0001$).

Table 4. Descriptive results for fluctuating asymmetry scores per population

Trait	Grote Kerk				Meerenberg			
	N	Mean	SD	Median	N	Mean	SD	Median
COBB	81	0.0189	0.0167	0.0200	38	0.0150	0.0118	0.0100
COBH	85	0.0208	0.0151	0.0200	36	0.0217	0.0148	0.0200
CNOR	84	0.0156	0.0125	0.0100	35	0.0177	0.0146	0.0100
CFMTN	89	0.0185	0.0136	0.0200	38	0.0145	0.0113	0.0100
CFMTNS	78	0.0169	0.0156	0.0100	33	0.0142	0.0123	0.0100
CMAH	86	0.0412	0.0339	0.0350	37	0.0478	0.0477	0.0400
CMPL	80	0.0648	0.0558	0.0500	36	0.0694	0.0541	0.0550
CMPB	84	0.0555	0.0462	0.0400	42	0.0643	0.0542	0.0500
CMSAST	61	0.0407	0.0349	0.0300	27	0.0370	0.0255	0.0300
COCL	99	0.0493	0.0462	0.0400	35	0.0463	0.0286	0.0500
COPO	100	0.0267	0.0238	0.0200	43	0.0265	0.0232	0.0200
CBAPO	107	0.0249	0.0216	0.0200	45	0.0222	0.0205	0.0200
CECMIS	30	0.0280	0.0225	0.0250	18	0.0328	0.0295	0.0250
CNMS	66	0.0179	0.0159	0.0200	30	0.0167	0.0132	0.0200
CBZO	75	0.0112	0.0097	0.0100	20	0.0125	0.0097	0.0150
CFMTB	86	0.0173	0.0138	0.0100	22	0.0177	0.0141	0.0200
CBPO	104	0.0194	0.0156	0.0200	24	0.0125	0.0126	0.0100
CBAST	86	0.0162	0.0138	0.0100	21	0.0181	0.0125	0.0200
CLFMT	79	0.0084	0.0074	0.0100	22	0.0136	0.0085	0.0100
CLAST	81	0.0221	0.0194	0.0200	19	0.0342	0.0315	0.0200
MAL	66	0.0136	0.0128	0.0100	28	0.0143	0.0126	0.0100
MRH	52	0.0394	0.0327	0.0300	27	0.0222	0.0241	0.0100
MXRB	67	0.0363	0.0273	0.0300	28	0.0282	0.0189	0.0250
MIRB	68	0.0406	0.0284	0.0400	40	0.0348	0.0256	0.0300
i1l_bl	21	0.0305	0.0267	0.0200	16	0.0163	0.0136	0.0100
i1l_md	17	0.0194	0.0164	0.0200	17	0.0159	0.0123	0.0100
i2l_bl	28	0.0243	0.0175	0.0200	16	0.0125	0.0068	0.0100
i2l_md	23	0.0252	0.0138	0.0200	18	0.0211	0.0164	0.0150
c1l_bl	30	0.0240	0.0189	0.0200	28	0.0261	0.0270	0.0200
c1l_md	30	0.0237	0.0214	0.0200	28	0.0218	0.0147	0.0200
pm1l_bl	39	0.0292	0.0190	0.0300	30	0.0297	0.0213	0.0300
pm1l_md	40	0.0383	0.0378	0.0200	31	0.0316	0.0268	0.0300
pm2l_bl	33	0.0285	0.0232	0.0200	20	0.0375	0.0271	0.0400
pm2l_md	33	0.0330	0.0325	0.0300	23	0.0435	0.0308	0.0400
m1l_bl	11	0.0236	0.0186	0.0300	9	0.0144	0.0101	0.0100
m1l_md	11	0.0218	0.0166	0.0100	10	0.0280	0.0210	0.0200
m2l_bl	20	0.0290	0.0229	0.0350	12	0.0125	0.0075	0.0100
m2l_md	24	0.0254	0.0245	0.0150	15	0.0333	0.0206	0.0300
i1u_bl	26	0.0196	0.0146	0.0200	18	0.0267	0.0209	0.0300
i1u_md	15	0.0200	0.0217	0.0100	9	0.0178	0.0172	0.0100

i2u_bl	19	0.0326	0.0249	0.0400	17	0.0318	0.0243	0.0200
i2u_md	10	0.0390	0.0277	0.0300	12	0.0450	0.0460	0.0300
c1u_bl	29	0.0241	0.0201	0.0200	26	0.0223	0.0203	0.0150
c1u_md	29	0.0210	0.0214	0.0100	25	0.0164	0.0125	0.0200
pm1u_bl	26	0.0162	0.0181	0.0100	15	0.0120	0.0086	0.0100
pm1u_md	25	0.0300	0.0214	0.0300	17	0.0459	0.0334	0.0400
pm2u_bl	28	0.0254	0.0228	0.0200	19	0.0179	0.0187	0.0100
pm2u_md	27	0.0467	0.0374	0.0500	17	0.0335	0.0255	0.0300
m1u_bl	12	0.0125	0.0087	0.0100	15	0.0087	0.0119	0.0000
m1u_md	13	0.0369	0.0180	0.0300	16	0.0438	0.0393	0.0350
m2u_bl	15	0.0253	0.0168	0.0200	19	0.0305	0.0255	0.0200
m2u_md	20	0.0495	0.0395	0.0450	19	0.0426	0.0323	0.0300
individual	134	0.0253	0.0071	0.0240	75	0.0261	0.0083	0.0239
dentition	64	0.0290	0.0118	0.0269	51	0.0272	0.0108	0.0250
cranium	121	0.0230	0.0072	0.0220	51	0.0235	0.0084	0.0236
c_orbit	83	0.0187	0.0104	0.0200	36	0.0192	0.0093	0.0200
c_facial	97	0.0211	0.0094	0.0200	41	0.0237	0.0178	0.0200
c_temporal	61	0.0407	0.0349	0.0300	27	0.0370	0.0255	0.0300
c_base	114	0.0297	0.0180	0.0275	48	0.0267	0.0154	0.0238
c_vault	109	0.0170	0.0115	0.0150	27	0.0175	0.0100	0.0150
mandible	77	0.0313	0.0159	0.0300	46	0.0289	0.0157	0.0258

N=number, SD = standard deviation, i=incisor, c=canine, pm=premolar, m=molar, l=lower, u=upper, bl=buccolingual, md=mesiodistal, c=cranium

Within the mandible, the minimum ramus breadth exhibited the highest mean FA, followed by the maximum ramus breadth, the height of the ramus, and mandibular length. Within each tooth class of both the mandibular and maxillary dentition, the mesiodistal crown diameters were more asymmetric than the corresponding buccolingual crown diameters. This pattern was reversed in the canines and the central incisors. Additionally, the level of asymmetry increased from mesial to distal within each tooth class, with the exception of the buccolingual diameter of the mandibular incisors, in which the second incisor exhibited less asymmetry than the first incisor.

The comparison of the frequency of pathological conditions between the populations, for the three lesions combined ($p=0.569$) and singly (Table 5), revealed no significant differences. However, relatively more GRK individuals exhibited cribra orbitalia and enamel hypoplasia (EH), while MeB individuals exhibited slightly more subperiosteal bone reactions (Table 5). Within both populations, subperiosteal bone reactions were observed most frequently (GRK=29.1%, MeB=30.7%), while the presence of cribra orbitalia was uncommon (GRK=6.7%, MeB=2.7%).

Table 5. Contingency table for each type of pathological condition observed in the individuals per population group.

Cribra orbitalia				
	Lesion (n)	No lesion (n)	Total	p-value
GRK	5	129	134	-
MeB	2	73	75	-
Total	7	202	209	1.00
Periosteal bone reactions				
	Lesion (n)	No lesion (n)	Total	p-value
GRK	39	95	134	-
MeB	23	52	75	-
Total	62	147	209	0.875
Enamel hypoplasia				
	Lesion (n)	No lesion (n)	Total	p-value
GRK	29	105	134	-
MeB	11	64	75	-
Total	40	169	209	0.272

When the median FA per individual who exhibited at least one of the three pathological conditions were compared between the populations, no significant differences were apparent ($p=0.480$), although the mean FA was slightly lower in the MeB population.

However, the comparison between the median FA of all individuals (index: Individual) who exhibited skeletal lesion(s) (2.6%; $\bar{x} = 2.7\%$, $\sigma = 0.7\%$) to individuals (index: Individual) without any of the three skeletal lesions (2.3%; $\bar{x} = 2.5\%$, $\sigma = 0.8\%$) revealed a significant difference ($p=0.009$). The median and mean FA values were greater in individuals with skeletal lesions.

4. DISCUSSION

The general trends observed in the cranium, mandible, and dentition concur with previous research suggesting that the magnitude of FA is variable across traits and indices (Blackburn, 2011; Møller and Swaddle, 1997; Storm, 2009). This differential expression of FA across traits can be attributed to the inherent differential canalization strength and ability of each trait (Waddington, 1957; Wagner et al., 1997). Each trait within an organism, or the same trait between organisms, can differ in their canalization strength against the same stressors. Additionally, traits can also vary in their response or buffering ability to differential stressors. That is to say, the same trait will react (buffer) differently to certain types of stressors than others (Møller and Swaddle, 1997; Waddington, 1957; Wagner et al., 1997). Similar to findings of other studies (Bigoni et al., 2013; e.g., DeLeon, 2007; Gawlikowska et al., 2007), FA was greatest in magnitude in the cranial base and lowest in the facial area. For

example, a study by DeLeon (2007) on Medieval Nubian crania found the vault or neurocranium to exhibit the greatest asymmetry, and the face the least asymmetry. Similarly, Medieval crania from central Europe exhibited the greatest amount of asymmetry in the cranial base, and the least in the upper facial area (Bigoni et al., 2013).

The magnitude of FA in the mandible was greatest for the minimum and maximum breadths, followed by the height of the ramus and mandibular length. This pattern is identical to findings by Storm (2009) on a large sample of British archaeological populations, but contradictory to previous findings by Storm and Knüsel (2005), where the breadth of the ramus exhibited the least asymmetry, and the ramus height the most. However, the latter study was conducted on a small subsample (n=20) of a British archaeological population, and might not be representative of the entire population. The mandible, as part of the craniofacial skeleton, responds or adapts to loading in a similar manner as the long bones, unevenly affecting the growth and bony response of the mandible (Kiliaridis, 2006). Components of the mandibular ramus, such as the condylar and coronoid processes, and the angle of the mandible, are especially affected by the muscular function of the mandible (Kiliaridis, 2006; Scott, 1953). Directional loading of the mandible due to habitual masticatory function (Pirttiniemi, 1998; Shah and Joshi, 1987), therefore, can be argued to be a contributing factor to increased FA levels of the breadths and height of the mandible. This might be indicative of the presence of DA within the mandible, despite that the statistical analyses did not show a significant presence of DA within any of the mandibular traits. Bilateral traits may exhibit both FA and DA simultaneously, and may have confounded or obscured the evaluation and expression of FA within the mandible (Møller and Swaddle, 1997; Van Valen, 1962).

The observed dental FA pattern is in accordance with the results of several other studies (Garn et al., 1966; Harris and Nweeia, 1980; Khalaf et al., 2005; Perzigian, 1977; Townsend and Brown, 1980), suggesting that dental crown formation is more stable for the mesial or first tooth in each morphogenic class, as well as for the buccolingual crown diameter of a tooth within a morphogenic class. This correlates with the general order of tooth and crown formation (most mesial tooth first), and with previous findings that the later forming teeth are more variable than the earlier forming teeth (Liversidge, 2003 and references within).

While both populations showed signs of cribra orbitalia, EH, and periostitis, these frequencies, as well as the magnitude of FA of these individuals were not significantly different between the populations. This may suggest that the populations were exposed to more or less the same magnitude or duration of external and internal stress during

development. However, it is also possible that the type of stress or the response to the stressors differed between the populations. For example, the alleged mental status of the MeB population might, besides the premise of increased developmental instability, also indicate that they were subjected to different types of stress relative to the GRK population. While various systemic and local factors can lead to EH, the hypoplastic defects sometimes occur with several syndromes and congenital illnesses that are also associated with mental or neurological disorders (Bhatia et al., 2012; Määttä et al., 2006; Schuurs, 2013; Webb et al., 1991). Therefore, if the observed higher frequencies of EH within the MeB population may, in part, be associated with their mental or neurological condition, the MeB population could be considered to have been subjected to less external stress, but to more internal stress relative to the GRK population. However, no genetic testing has been attempted, and therefore, all possible etiologies, including nutritional deficiencies and trauma, should be considered before such a conclusion can be reached. In addition, the scoring of EH was not comprehensive, limiting the amount of information which could have been valuable for assessment, such as the time (stage of development) at which the insults occurred within each population.

In line with the abovementioned lack of a difference in terms of pathology and in contrast to findings in the literature that individuals with a history of mental disorders or neurological impairments exhibit higher levels of FA (Lalumière et al., 2001; Malina and Buschang, 1984; Reilly et al., 2001), no significant difference in the FA values was apparent between the GRK and MeB populations. Mentally ill individuals have been shown to be more developmentally unstable, and are therefore expected to exhibit higher levels of FA compared to their mentally healthy counterparts (Lalumière et al., 2001; Malina and Buschang, 1984; Martin et al., 1999; Reilly et al., 2001). For example, Malina and Buschang (1984), observed higher levels of FA within the long bones of mentally institutionalized males compared to mentally healthy males, which were statistically significant for the upper limb traits.

While this lack of a difference in FA magnitude between the two populations may be an indication that FA is not such a highly sensitive indicator of stress as indicated in the literature (DeLeon, 2007; Hoover and Matsumura, 2008; Livshits and Kobylansky, 1991; Storm, 2008, 2007), it is important to also interpret the results in relation to the context in which the individuals lived. Comparisons of the socio-political environments of the individuals from GRK during the 18th to early 19th century, and individuals from MeB during the early 19th to the early 20th centuries, suggest that the GRK population experienced a more turbulent economic environment (Allen, 2001, 2000; Deneweth et al., 2014; De Vries and

Van der Woude, 1997; Haines, 2004; Van den Eerenbeemt, 1962). The 18th to early 19th century (GRK population) Dutch economy started and ended amidst the remnants of economic recessions with a period of slow economic growth in between (1740-1780). Rather turbulent in its course, the average standard of living decreased over the entire 18th century, reaching an ultimate low at the turn of the century. Poor living standards resulted from a combination of the stagnation of wage increases and an increase in the cost of living. While the economic recession ended at the turn of the century, it left a nation in poverty (Allen, 2000; Deneweth et al., 2014; De Vries, 1968; De Vries and Van der Woude, 1997). Despite the adverse economic conditions between 1795 and 1813, the early 19th century (GRK and MeB populations) saw a rise in agriculture and the overall economy, which was followed by rapid rates of industrialization and wage increases during the late 19th century (MeB population).

While not yet stable during the 19th century, the economy seemed to have steadily improved during the time in which the MeB population lived. The national earnings per capita showed an overall, albeit turbulent, increase throughout the 19th and early 20th century, while a decline in the average stature and an impoverished standard of living was evident from the 1830's to 1860's (Allen, 2001; Haines, 2004). However, from about 1870 onward, the country experienced rapid population growth together with an increase in adult stature, and an increased life expectancy and standard of living (Haines, 2004; Maat, 2005). The standard of living of Dutch laborers continued to improve into the early 20th century (Deneweth et al., 2014).

When the information on the pathological lesions and the socio-political history of the two populations are considered simultaneously, it is possible that the GRK population was subjected to greater amounts (or longer duration) of external stressors, while the MeB population was subjected to less external stress, but also to internal stress (due to their mental status). However, the similar frequencies of pathological lesions between GRK and MeB can also be an indication that despite the socio-economic history, both populations were subjected to similar amounts or durations of external stress.

However, it should also be considered that the MeB psychiatric hospital sample was not as developmentally unstable as suggested by previous research (Malina and Buschang, 1984; Martin et al., 1999; Reilly et al., 2001). Unfortunately, a direct comparison of FA magnitude between the Dutch populations and other populations (with and without mental deficiencies) reported in the literature is difficult due to the effect of trait size on FA estimates. The magnitude of FA is influenced not only by the interaction of trait size, but FA

may also be proportional to the size of a trait (Palmer and Strobeck, 2003). Therefore, if the effect of size was not accounted for in a study, differences in trait size between studies would differentially have affected the magnitude of FA, inflating or decreasing the real difference in asymmetry. Consequently, a comparison of FA magnitude between the two Dutch populations and previous studies on the mentally institutionalized was not attempted. However, this study calculated FA estimates in a similar manner and for identical cranial and mandibular traits as a study by Storm (2007, 2008) on English archaeological populations, allowing for a comparison of the FA magnitudes. When compared to each other, the average adult FA for the cranial and mandibular traits in the Dutch populations are greater than that of the English populations, although the FA differences are similar in magnitude compared to the difference in average FA between the two Dutch populations.

While no significant differences were noted in the FA magnitude between the populations and despite the fact that the pathological lesions were assessed in a very basic manner, evidence supporting the use of FA as an indicator of developmental instability and its positive relationship to skeletal indicators of stress was observed in the comparison of the magnitude of FA between individuals exhibiting a pathological condition and the rest of the sample. Individuals who exhibited a pathological condition were significantly more asymmetrical (in magnitude) than the individuals without any of the three conditions. This corresponds to findings by Hoover and Matsumura (2008) on 13 archaeological Japanese populations, in which higher levels of FA were observed in individuals who exhibited EH. It can be inferred that individuals who are better buffered or resistant against environmental stressors, such as nutritional deficiencies or infectious diseases, are also more developmentally stable.

However, fluctuating asymmetry is a non-specific indicator of stress, and therefore the average FA in a population does not distinguish between the type of stress (Møller and Swaddle, 1997; Palmer, 1994). This is also relevant in the interpretation of the apparent similar FA magnitudes between the two Dutch populations. Without context the lack of a difference in FA between the two populations could be interpreted to be indicative of differential levels of developmental instability under different types and durations of stress, such as a more unstable development with less severe or shorter durations of stress for the MeB population, and a more stable developmental with severe or longer durations of stress for the GRK population. However, the observed patterns and similar magnitudes of FA, in addition to their socio-economic background and frequencies of indicators of pathology or

stress, indicate that both populations were likely subjected to similar levels of stress, even though the source, timing or duration of stress might have been different.

Another consideration is that a significant difference in their developmental stability does in fact exist, but it was undetectable using these methods. Factors such as sample size and measurement error may affect the detection of departures from FA within a sample. Small sample sizes, for example, may obscure the magnitude of FA in such a way that no FA is detected, even though a departure from FA may exist (Palmer, 1994; Storm, 2009). While the sample sizes for both populations were above the desirable sample size of 40 or 50, as suggested by Palmer (1994), the GRK sample was nearly two times larger than the MeB sample. In addition, when categorized per sex and age category, the sample sizes for each category were no longer above that of the minimum of 30. Consequently, these differences in sample size, especially the smaller sample sizes, could possibly have obscured from the detection of FA. What was observed as a lack of statistical differences in FA between the age categories, the sexes, and the populations within this study could, therefore, have been influenced or obscured by the differences in sample size.

As with all studies on the deceased, it is important to consider that the individuals under study are those who did not survive. Known as hidden heterogeneity, individuals within a population are differentially susceptible to various diseases and stressors, which is unknown to the researcher (DeWitte and Stojanowski, 2015; Wood et al., 1992). Applied to the current study, it can be argued that the observed levels of FA are elevated against what would have been the FA of the entire population. An increased magnitude of FA within a population has been associated with decreased fitness of populations (Clarke, 1995). Therefore, it can be inferred that the individuals in the sample with a higher magnitude of FA, and a lower level of fitness, would have had an increased risk of dying – increasing the average magnitude of FA within the skeletal sample above that of the entire population.

5. CONCLUSIONS

The main aim of this study was to assess the developmental stability between two Dutch populations, the MeB psychiatric hospital sample and the GRK general population sample by means of comparing the magnitude and patterns of FA within the cranium, mandible, and dentition. The main conclusions are as follows:

- The magnitude of FA is variable across traits and indices, which can be linked to their differential buffering to differential types of stressors.

- The two populations exhibited similar frequencies of pathological lesions, as well as similar levels of FA between the individuals exhibiting these lesions.
- In contrast to findings in the literature, the MeB psychiatric hospital sample was similarly asymmetric compared to the general population of the GRK (Alkmaar). The socio-economic history and observed frequencies of pathological lesions suggest that both populations were likely subjected to similar levels of stress, even though the source, timing, and duration of stress might have been different. Another possibility is that the MeB population was not as developmentally unstable as suggested by previous research.
- When the samples were combined the individuals with signs of at least one of the three pathological lesions exhibited significantly greater FA values compared to the rest of the sample. This positive relationship between FA to skeletal and dental indicators of stress support the use of FA as an indicator of developmental instability.

6. ACKNOWLEDGEMENTS

The financial assistance of the National Research Foundation (NRF) of South Africa towards this research (M Steyn and A Hagg) is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the authors and are not necessarily those of the NRF. Special thanks go to Clarisa Sutherland for her assistance in this research project, and to Marinda Pretorius for her joyful work with the illustrations included within this article.

7. LITERATURE CITED

- Acsádi, G., & Nemeskéri, J. (1970). *History of human life span and mortality*. Budapest: Akadémiai Kiadó.
- Adams, M., & Niswander, J. (1967). Developmental “noise” and a congenital malformation. *Genet Res*, 10, 313–317.
- Allen, R. (2001). The great divergence in European wages and prices from the Middle Ages to the First World War. *Explor Econ Hist*, 38, 411–447.
- Allen, R. (2000). Economic structure and agricultural productivity in Europe, 1300-1800. *Eur Rev Econ Hist*, 3, 1–25.
- Arnqvist, G., & Thornhill, R. (1998). Evolution of animal genitalia: Patterns of phenotypic and genotypic variation and condition dependence of genital and non-genital morphology in water strider (Heteroptera: Gerridae: Insecta). *Genet Res*, 71, 193–212.
- Baetsen, S. (2001). *Graven in de Grote Kerk. Het fysich-antropologisch onderzoek van de graven in de St. Laurenskerk van Alkmaar*. Alkmaar: Gemeente Alkmaar.
- Baetsen, S., Bitter, P., & Bruintjes, T. (1997). Hip and knee osteoarthritis in an eighteenth century urban population. *Int J Osteoarchaeol*, 7, 628–630.
- Bai, J., & Ng, S. (2005). Tests for skewness, kurtosis, and normality for time series data. *J Bus Econ Stat*, 23(1), 49–60.

- Bailit, H., Workman, P., Niswander, J., & Lean, C. (1970). Dental asymmetry as an indicator of genetic and environmental conditions in human populations. *Hum Biol*, 42(4), 626–638.
- Bhatia, S., Dubey, G., Kapur, A., & Ritwik, P. (2012). Congenital rubella syndrome: Dental manifestations and management in a 5 year old child (abstract only). *J Clin Pediatr Dent* 37(1), 71–75.
- Bigoni, L., Krajčček, V., Sládek, V., Velemínský, P., & Velemínská, J. (2013). Skull shape asymmetry and the socioeconomic structure of an early Medieval central European society. *Am J Phys Anthropol*, 150, 349–364.
- Bitter, P. (2002). *Graven en begraven. Archeologie en geschiedenis van de Grote Kerk van Alkmaar*. University of Amsterdam, Hilversum.
- Bjorksten, T., David, P., Pomiankowski, A., & Fowler, K. (2000). Fluctuating asymmetry of sexual and nonsexual traits in stalk-eyed flies: A poor indicator of developmental stress and genetic quality. *J Evol Biol*, 13, 89–97.
- Blackburn, A. (2011). Bilateral asymmetry of the humerus during growth and development. *Am J Phys Anthropol*, 145(4), 639–646.
- Black, T. (1980). An exception to the apparent relationship between stress and fluctuating dental asymmetry. *J Dent Res*, 59(7), 1168–1169.
- Brickley, M., Ives, R., 2008. *The bioarchaeology of metabolic bone disease*, 1st ed. Elsevier Ltd., Hungary.
- Brook, A. (2009). Multilevel complex interactions between genetic, epigenetic and environmental factors in the aetiology of anomalies of dental development. *Arch Oral Biol*, 54s, s3–s17.
- Brooks, S., & Suchey, J. (1990). Skeletal age determination based on the os pubis: A comparison of the Acsadi-Nemeskeri and Suchey-Brooks methods. *J Hum Evol*, 5(3), 227–238.
- Buikstra, J., & Ubelaker, D. (Eds.). (1994). *Standards for data collection from human skeletal remains*. Arkansas: Arkansas Archeological Survey.
- Clarke, G. (1995). Relationships between developmental stability and fitness: Application for conservation biology. *Conserv Biol*, 9(1), 18–24.
- Dahlberg, G. (1940). *Statistical methods for medical and biological students*. London: George Allen & Unwin Ltd.
- DeLeon, V. (2007). Fluctuating asymmetry and stress in a Medieval Nubian population. *Am J Phys Anthropol*, 132, 520–534.
- Deneweth, H., Gelderblom, O., & Jonker, J. (2014). Microfinance and the decline of poverty: Evidence from the nineteenth-century Netherlands. *J Econ Dev*, 39(1), 79–110.
- De Vries, J. (1968). Economische groei en industrialisatie in Nederland 1850-1914. *Maandschr Econ*, 33(3), 118–128.
- De Vries, J., & Van der Woude, A. (1997). *The first modern economy: Success, failure, and perseverance of the Dutch economy, 1500-1815*. Cambridge: Cambridge University Press.
- DeWitte, S., & Stojanowski, C. (2015). The Osteological Paradox 20 years later: Past perspectives, future directions. *J Archaeol Res*, 23, 397–450.
- Falys, C., & Lewis, M. (2011). Proposing a way forward: A review of standardisation in the use of age categories and aging techniques in osteological analysis (2004-2009). *Int J Osteoarchaeol*, 21(6), 704–716.
- Fields, S., Spiers, M., Hershkovitz, I., & Livshits, G. (1995). Reliability of reliability coefficients in the estimation of asymmetry. *Am J Phys Anthropol*, 96, 83–87.
- Garn, S., Lewis, A., & Kerewsky, R. (1966). The meaning of bilateral asymmetry in the permanent dentition. *Angle Orthod*, 36(1), 55–62.

- Gawlikowska, A., Szczurowski, J., Czerwiński, F., Miklaszewska, D., Adamiec, E., & Dzieciołowska, E. (2007). The fluctuating asymmetry of mediaeval and modern human skulls. *J Comp Hum Biol*, 58, 159–172.
- Geeta, A., Jamaiyah, H., Safiza, M., Khor, G., Ahmad, A., Suzana, S., ... Faudzi, A. (2009). Reliability, technical error of measurements and validity of instruments for nutritional status assessment of adults in Malaysia. *Singapore Med J*, 50(10), 1013–1018.
- Goldberg, D., & Graham, T. (2013). *Mental health in our future cities*. Psychology Press.
- Grubbs, F. (1969). Procedures for detecting outlying observations in samples. *Technometrics*, 11(1), 1–21.
- Guatelli-Steinberg, D., Sciulli, P., & Edgar, H. (2006). Dental fluctuating asymmetry in the Gullah: Tests of hypotheses regarding developmental stability in deciduous vs. permanent and male vs. female teeth. *Am J Phys Anthropol*, 129, 427–434.
- Haines, M. (2004). Growing incomes, shrinking people - Can economic development be hazardous to your health? Historical evidence for the United States, England, and the Netherlands in the nineteenth century. *Soc Sci Hist*, 28(2), 249–270.
- Harris, E., & Nweeia, M. (1980). Dental asymmetry as a measure of environmental stress in the Ticuna Indians of Columbia. *Am J Phys Anthropol*, 53, 133–142.
- Hartnett, K. (2010). Analysis of age-at-death estimation using data from a new, modern autopsy sample-Part 1: Pubic bone. *J Forensic Sci*, 55(5).
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scand J Stat*, 6(2), 65–70.
- Hoover, K., Corruccini, R., Bondioli, L., & Macchiarelli, R. (2005). Exploring the relationship between hypoplasias and odontometric asymmetry in Isola Sacra, an imperial Roman Necropolis. *Am J Hum Biol*, 17, 752–764.
- Hoover, K., & Matsumura, H. (2008). Temporal variation and interaction between nutritional and developmental instability in prehistoric Japanese populations. *Am J Phys Anthropol*, 137, 469–478.
- Howells, W. (1973). *Cranial variation in man: A study by multivariate of patterns of differences among recent human populations*. Cambridge: Harvard University Press.
- İşcan, M., Loth, S., & Wright, R. (1985). Age estimation from the rib by phase analysis: White females. *J Forensic Sci*, 30, 853–863.
- İşcan, M., Loth, S., & Wright, R. (1984a). Metamorphosis at the sternal rib: A visual assessment technique. *Am J Phys Anthropol*, 65, 147–156.
- İşcan, M., Loth, & Wright, R. (1984b). Age estimation from the rib by phase analysis: White males. *J Forensic Sci*, 29, 1094–1104.
- Khalaf, K., Elcock, C., Smith, R., & Brook, A. (2005). Fluctuating dental asymmetry of multiple crown variables measured by an image analysis system. *Arch Oral Biol*, 50, 249–253.
- Kieser, J. (1990). *Human adult odontometrics*. Cambridge: Cambridge University Press.
- Kiliaridis, S. (2006). The importance of masticatory muscle function in dentofacial growth. *Semin Orthod*, 12(2), 110–119.
- Klales, A., Ousley, S., & Vollner, J. (2012). A revised method of sexing the human innominate using Phenice's nonmetric traits and statistical methods. *Am J Phys Anthropol*, 149, 104–114.
- Lalumière, M., Harris, G., & Rice, M. (2001). Psychopathy and developmental instability. *Evol Hum Behav*, 22, 75–92.
- Laurikkala, J., Juhola, M., & Kentala, E. (2000). *Informal identification of outliers in medical data*. Paper presented at the fifth international workshop on Intelligent Data Analysis in Medicine and Pharmacology, Berlin, Germany.

- Liversidge, H. (2003). Variation in modern human dental development. In J. Thompson, G. Krovitz, & A. Nelson (Eds.), *Patterns of growth and development in the genus Homo* (pp. 73–113). New York: Cambridge University Press.
- Livshits, G., & Kobylansky, E. (1991). Fluctuating asymmetry as a possible measure of developmental homeostasis in humans: A review. *Hum Biol*, 63(4), 441–466.
- Maat, G. (2005). Two millennia of male stature development and population health and wealth in the Low Countries. *Int J Osteoarchaeol*, 15, 276–290.
- Määttä, T., Tervo-Määttä, T., Taanila, A., Kaski, M., & Iivanainen, M. (2006). Mental health, behaviour and intellectual abilities of people with Down Syndrome. *Down Syndr Res Pract*, 11(1), 37–43.
- Malina, R., & Buschang, P. (1984). Anthropometric asymmetry in normal and mentally retarded males. *Ann Hum Biol*, 11(6), 515–531.
- Mann, R., & Murphy, S. (1990). *Regional atlas of bone disease: A guide to pathologic and normal variation in the human skeleton*. Illinois: Charles C Thomas.
- Markow, T., & Gottesman, I. (1989). Fluctuating dermatoglyphic asymmetry in psychotic twins. *Psychiatry Res*, 29, 37–43.
- Markow, T., & Wandler, K. (1986). Fluctuating dermatoglyphic asymmetry and the genetics of liability to schizophrenia. *Psychiatry Res*, 19, 323–328.
- Martin, S., Manning, J., & Dowrick, C. (1999). Fluctuating asymmetry, relative digit length, and depression in men. *Evol Hum Behav*, 20, 203–214.
- Massey, F. (1951). The Kolmogorov-Smirnov test for goodness of fit. *J Amer Statist Assoc*, 46(253), 68–78.
- Meindl, R., & Lovejoy, C. (1985). Ectocranial suture closure: A revised method for the determination of skeletal age at death based on the lateral-anterior sutures. *Am J Phys Anthropol*, 68, 57–66.
- Møller, A., & Swaddle, J. (1997). *Asymmetry, developmental stability, and evolution*. (R. May & P. Harvey, Eds.). Oxford: Oxford University Press.
- Morrison, D. (1969). On the interpretation of discriminant analysis. *J Mark Res*, 6(2), 156–163.
- Norton, K., & Olds, T. (Eds.). (1996). *Anthropometrica: A textbook of body measurement for sports and health courses* (First). Sydney: University of New South Wales Press.
- Ortner, D. (2003). *Identification of pathological conditions in human skeletal remains*. London: Academic Press.
- Palmer, A. (1994). Fluctuating asymmetry analyses: A primer. In T. Markow (Ed.), *Developmental instability: Its origins and evolutionary implications* (pp. 335–364). Dordrecht: Kluwer.
- Palmer, A., & Strobeck, C. (2003). Fluctuating asymmetry analyses revisited. In M. Polak (Ed.), *Developmental Instability: Causes and Consequences* (pp. 279–319). Oxford University Press.
- Palmer, A., & Strobeck, C. (1986). Fluctuating asymmetry: Measurement, analysis, patterns. *Annu Rev Ecol Syst*, 17, 391–421.
- Perini, T., de Oliveira, G., dos Santos Ornellos, J., & de Oliveira, F. (2005). Technical error of measurement in anthropometry. *Rev Bras Med Esporte*, 11(1), 86–90.
- Perzigian, A. (1977). Fluctuating dental asymmetry: Variation among skeletal populations. *Am J Phys Anthropol*, 47, 81–88.
- Phenice, T. (1969). A newly developed visual method of sexing the os pubis. *Am J Phys Anthropol*, 30(2), 297–301.
- Pirttiniemi, P. (1998). Normal and increased functional asymmetries in the craniofacial area. *Acta Odontol Scand*, 56(6), 342–345.

- Rao, J., & Scott, A. (1984). On chi-squared tests for multiway contingency tables with cell proportions estimated from survey data. *Ann Stat*, 12(1), 46–60.
- Reilly, J., Murphy, P., Byrne, M., Larkin, C., Gill, M., O’Callaghan, E., & Lane, A. (2001). Dermatoglyphic fluctuating asymmetry and atypical handedness in schizophrenia. *Schizophr Res*, 2001, 159–168.
- Reitsema, L., & McIlvaine, B. (2014). Reconciling “stress” and “health” in physical anthropology: What can bioarchaeologists learn from the other subdisciplines? *Am J Phys Anthropol*, 155(2), 181–185.
- Schuurs, A. (2013). *Pathology of the hard dental tissues*. West Sussex: John Wiley & Sons.
- Scott, J. (1953). The growth of the human face. In *Section of odontology* (Vol. 47, pp. 91–100).
- Seow, W. (1997). Effects of preterm birth on oral growth and development. *Aust Dent J*, 42(2), 85–91.
- Shackelford, T., & Larsen, R. (1997). Facial asymmetry as an indicator of psychological, emotional, and physiological distress. *J Pers Soc Psychol*, 72(2), 456–466.
- Shah, S., & Joshi, M. (1987). An assessment of asymmetry in the normal craniofacial complex. *Angle Orthod*, 48(2), 141–148.
- Storm, R. (2009). *Human skeletal asymmetry: A study of directional and fluctuating asymmetry in assessing health, environmental conditions, and social status in English populations from the 7th to the 19th centuries* (Doctoral thesis). University of Bradford, United Kingdom.
- Storm, R. (2008). Cranial asymmetry and developmental abnormalities. In J. Magilton, F. Lee, & A. Boylston, *Lepers outside the gate: Excavations at the cemetery of the hospital of St James and St Mary Magdalene, Chichester, 1986-87 and 1993* (1st ed., p. 294). York: Council for British Archaeology.
- Storm, R. (2007). *The stressful revolution: A rise in fluctuating asymmetry from Medieval to Victorian England*. Paper presented at the proceedings of the seventh annual conference of the British Association for Biological Anthropology and Osteoarchaeology (pp. 95–104), London, England.
- Storm, R., & Knüsel, C. (2005). *Fluctuating asymmetry: A potential osteological application*. Paper presented at proceedings of the fifth annual conference of the British Association for Biological Anthropology and Osteoarchaeology, Southampton, England.
- Townsend, G., & Brown, T. (1980). Dental asymmetry in Australian Aborigines. *Hum Biol*, 52(4), 661–673.
- Ulijaszek, S., & Kerr, D. (1999). Anthropometric measurement error and the assessment of nutritional status. *Br J Nutr*, 82, 165–177.
- Van den Eerenbeemt, H. (1962). Oorzaken van het pauperisme in Nederland in de 18e eeuw. *Maandschr Econ*, 27(1-2), 156–166.
- Van der Merwe, A., Weston, D., Oostra, R., & Maat, G. (2013). A review of the embryological development and associated developmental abnormalities of the sternum in the light of a rare palaeopathological case of sternal clefting. *J Comp Hum Biol*, 64, 129–141.
- Van Twuyver, P. (2000). *Meerenberg 150 jaar: Meer dan een gesticht: Een historie in foto’s*. Haarlem: Stichting GGZ museum, Pest- en Dolhuys.
- Van Valen, L. (1962). A study of fluctuating asymmetry. *Int J Org Evol*, 16(2), 125–142.
- Vijsselaar, J. (1982). *Krankzinnigen gesticht: Psychiatrische inrichtingen in Nederland 1880 - 1910*. Bussum: Unieboek.
- Waddington, C. (1957). Chapter 2. The cybernetics of development. In *The strategy of the genes: A discussion of some aspects of theoretical biology*. London: George Allen & Unwin Ltd.

- Waddington, C. (1942). Canalization of development and the inheritance of acquired characters. *Nature*, 150(3811), 563–565.
- Wagner, G., Booth, G., & Baghari-Chaichian, H. (1997). A population genetic theory of canalization. *Evolution*, 51(2), 329–347.
- Waldron, T. (2009). *Palaeopathology*. Cambridge: Cambridge University Press.
- Webb, D., Fryer, A., & Osborne, J. (1991). On the incidence of fits and mental retardation in tuberous sclerosis. *J Med Genet*, 28, 395–397.
- Weisensee, K. (2013). Assessing the relationship between fluctuating asymmetry and cause of death in skeletal remains: A test of the developmental origins of health and disease hypothesis. *Am J Hum Biol*, 25, 411–417.
- Wood, J., Milner, G., Harpending, H., & Weiss, K. (1992). The Osteological Paradox: Problems of inferring prehistoric health from skeletal samples. *Curr Anthropol*, 33(4), 343–370.