Monitoring of Therapy for Mucopolysaccharidosis Type I Using Dymorphometric Facial Phenotypic Signatures

Stefanie Kung • Mark Walters • Peter Claes •
Peter LeSouef • Jack Goldblatt • Andrew Martin •
Shanti Balasubramaniam • Gareth Baynam

Received: 24 November 2014 / Revised: 29 January 2015 / Accepted: 30 January 2015 / Published online: 03 March 2015
© SSIEM and Springer-Verlag Berlin Heidelberg 2015

Abstract There is a pattern of progressive facial dysmorphology in mucopolysaccharidosis type I (MPS I). Advances in 3D facial imaging have facilitated the development of tools, including dymorphometrics, to objectively and

precisely detect these facial phenotypes. Therefore, we investigated the application of dymorphometrics as a noninvasive therapy-monitoring tool, by longitudinally scoring facial dysmorphology in a child with MPS I receiving enzyme replacement therapy (ERT) and bone marrow transplantation (BMT). Both dymorphometric measures showed a decreasing trend, and the greatest differences were found in the severity of facial discordance (Z-RMSE), displaying scores >3 SD higher than the mean at their peak, in comparison to Z-RSD scores that mostly fell within the normative range (maximum; 1.5 SD from the mean). In addition to the general trend of reduced facial dysmorphology with treatment, initial fluctuations were also evident that may have related to transient subcutaneous facial fluctuations, in the context of conditioning for bone marrow transplant. These findings support the potential of our approach as a sensitive, noninvasive, and rapid means of assessing treatment response or failure in clinical trials, and for established therapies, and would be applicable for other inherited disorders of metabolism.

Communicated by: Frits Wijburg, MD, PhD

Competing interests: None declared

S. Kung (✉) • P. LeSouef • J. Goldblatt • A. Martin • G. Baynam
School of Paediatrics and Child Health, University of Western
Australia, GPO Box D184, Perth, WA 6840, Australia
e-mail: Stefanie.Kung@health.wa.gov.au

M. Walters
Cranio-maxillo-facial Unit, Princess Margaret Hospital for Children,
Perth, WA, Australia

P. Claes
Department of Electrical Engineering, ESAT/PSA, KU Leuven,
Medical Image Computing, Leuven, Belgium

P. Claes
UZ Leuven, Medical Imaging Research Center, Leuven, Belgium

P. Claes
iMinds, KU Leuven, Medical IT Department, Leuven, Belgium

J. Goldblatt • G. Baynam (✉)
Genetic Services of Western Australia, King Edward Memorial &
Princess Margaret Hospitals, 4th Floor Agnes Walsh House, 374
Bagot Road, Subiaco, WA 6008, Australia
e-mail: Gareth.Baynam@health.wa.gov.au

J. Goldblatt • G. Baynam
Office of Population Health Genomics, Western Australian
Department of Health, Perth, WA, Australia

S. Balasubramaniam
Storage Disorders Clinic, Princess Margaret Hospital for Children,
Perth, WA, Australia

G. Baynam
Institute for Immunology and Infectious Diseases, Murdoch
University, Murdoch, WA, Australia

Abbreviations

3D	Three dimensional
AM	Anthropometric mask
BMT	Bone marrow transplant
ERT	Enzyme replacement therapy
GvHD	Graft-versus-host disease
HSCT	Hematopoietic stem cell transplantation
LSD	Lysosomal storage disorder
MPS IH	Mucopolysaccharidosis type I – Hurler syndrome
NE	Normal equivalent
PCA	Principal component analysis
PCs	Principal components
RMSE	Root mean squared error

RSD Relative significant discordance
SD Standard deviation

Introduction

Lysosomal storage disorders (LSD), which include mucopolysaccharidosis type I (MPS I), are a heterogeneous group of genetic disorders caused by a deficiency of one or more degradation enzymes essential for normal cell metabolism (Muhlstein et al. 2013). The lack of α -L-iduronidase leads to multisystemic accumulation of its substrates within the lysosomes. This causes a variably expressed systemic disorder, which can be clinically classified into severe or attenuated forms; this distinction influences therapeutic options (Clarke and Heppner 2011).

The development of MPS I-targeted treatments, including enzyme replacement therapy (ERT) and allogeneic hematopoietic stem cell transplantation (HSCT), has dramatically changed prognosis. However, limitations of current treatments and cost of therapy are motivating the development of novel approaches (van Gelder et al. 2012). This requires the means to objectively and recurrently assess treatment response. These assessments will preferentially be noninvasive, deeply precise, relatively inexpensive, and portable. Existing monitoring modalities are limited in their ability to objectively document responses to therapy following short-term clinical trials due to the variability of the phenotypes, the irreversibility of some complications, and the invasive nature of some investigations. Assessing response to therapy for MPS I has generally included physical/mobility tests to examine joint function, the 6-min walking test (6MWT), lung function, changes in hepatosplenomegaly, and biochemical assays of glycosaminoglycan (GAG) substrate (Church et al. 2007).

MPS I has a pattern of progressive facial dysmorphology, particularly in untreated cases. Advances in 3D facial imaging have facilitated the development of anthropometric tools, including dysmorphometrics, to objectively detect these facial phenotypes (Claes et al. 2012a, b, 2013; Hammond and Suttie 2012). Previously, dysmorphometrics was used to cross-sectionally detect and localize MPS I-associated facial dysmorphologies and, thereby, establish an objective facial phenotype. These individual signatures were attributed severity scores to discriminate between individuals clinically diagnosed with MPS I subtypes (Kung et al. 2012). In this study, we investigate the application of dysmorphometrics as a noninvasive treatment-monitoring tool, by longitudinally scoring facial dysmorphology in a treated MPS I-affected child.

Methods

Participants

A normative reference cohort of approximately 1,000 individuals was obtained from the Perth Face-Space Project, aged 1 month to 25 years and of self-reported ancestry. Participants completed a questionnaire on relevant medical history, and those with prior craniofacial surgery or a suspected syndromic condition with craniofacial manifestations were excluded from this cohort. This reference cohort was imperative for the construction of patient-specific controls. Ethics approvals (PMHEC: 1801/EP, 1443/EP, and 1488/EP) were granted by Princess Margaret Hospital for Children Ethics Committee in Perth, Western Australia.

The child with severe MPS I was recruited from the Princess Margaret Hospital for Children in Perth, Western Australia. He initially presented to the Emergency Department at 6 months of age with a cough, and a chest X-ray revealed paddle-shaped ribs, suggestive of a mucopolysaccharidosis. Weekly ERT infusions began at 10 months of age, followed by a bone marrow transplantation at 12 months of age, with a further 3 months of ERT post-BMT. 3D facial scans were ascertained at eight time points.

3D Image Acquisition, Data Preparation

3D facial scans were captured using a 3dMDFacial™ stereophotogrammetric camera system, and facial shape was expressed as a point cloud of approximately 200,000 points in a 3D coordinate space, the reliability and precision of which have been validated (Aldridge et al. 2005). These facial scans or point cloud data were brought into closer alignment by manual indication of 12 anatomical landmarks (right/left exocanthion, right/left endocanthion, pronasale, lateral nasal ala corners, right/left cheilion, upper lip tubercle, vermillion border, and chin point), in preparation for the facial mapping process. This decreases computational time and provides a basis for surface registration during facial mapping.

Anthropometric Masks and Facial Mapping

A spatially dense indicated set of 10,000 quasi-landmarks was obtained using a nonrigid surface registration (mapping) technique of a predefined facial template (anthropometric mask) (Claes et al. 2012b). This fully automated facial mapping process was performed across all faces in the dataset and required re-sampling of the raw point cloud data into a more comparable format. A reference scan, known as an anthropometric mask and representing the standard of connectivity for all scans, was

then created through an iterative “bootstrapping” method, as described by Claes (2007). Once facial scans were mapped, dense anatomical correspondence was achieved and allowed for biologically valid comparisons to be performed.

Reference Face Space

To define the statistical limits of typical facial variation in a normative reference population, a statistical face space was constructed from 1,000 individuals in our reference range cohort. A generalized Procrustes fit rotated, translated, and scaled the quasi-landmark configurations into the same coordinate space, where shape variation was described by Procrustes distance residuals. This statistical face space, via principal component analysis (PCA), describes variations in facial form and elucidates complex harmonic interrelationships between these variations.

Dysmorphometrics

Using the dysmorphometric approach (Claes et al. 2012a), normative references were encoded within the face space, and outliers in comparison to this normative reference reflected discordancy in the facial form (i.e., facial dysmorphology). This process involves the robust superimposition of the reference face space onto the patient’s facial scan, where each of the 10,000 quasi-landmarks is assigned a confidence value against a p -value of 0.05. A value closer to 1 reflects the tendency of that point being harmonic (inlier), while a value closer to 0 reflects its tendency of being discordant (outlier). Through this superimposition, patient-specific and population-based matched references called “normal equivalents” (NE) were generated. The NE (Claes et al. 2013) describes any given face in terms of harmonious facial variation (i.e., patient’s facial scan without the dysmorphology), and its construction was performed without any a priori knowledge of the condition itself. NE facial scans were generated at the eight assessment time points during the treatment course.

Normal Equivalent Facial Assessments

To analyze the facial discordancy at each time point, each NE was superimposed onto its chronologically corresponding patient scan. The differences between corresponding landmarks of each NE patient scan pair provided the means to measure both the magnitude and vectors (direction) of the observed facial discordances. Global scores RSD (relative significant discordance; %) and RMSE (root mean squared error; mm) provide an

overall measure of discordant facial proportions and discordance severity, respectively. The NE assessment also outputs two dysmorphograms that enable visualization of discordances on a facial manifold, namely, (1) distance and (2) outlier facial maps. Distance facial maps highlight localized regions of RMSE, which take into account both variance and possible bias, as an error in millimeters (mm). Outlier (confidence) maps highlight localized regions of RSD on the facial surface, while vector maps provide directional information on the observed facial discordance. Collectively, these distance, outlier, and vector maps provide an individualized dysmorphometric signature. The method workflow presented in this study is summarized in the Fig. 1. Normalized Z-scores (Z-RSD, Z-RMSE) were also generated from reference summary NE statistics obtained from the normative population, RSD (mean 10.6 and SD 1.8) and RMSE (mean 0.91 and SD 0.22).

Results

The objective measures showed similar trends for both global RSD and RMSE discordance scores over the treatment course; their regional differences are highlighted in facial outlier and distance maps, respectively (Fig. 2). The greatest amount of change was seen at the lower two-thirds of the face; in particular, facial discordance at the nasal, perioral/labial, cheek, and mandibular regions diminished over the treatment course. The fullness of the upper lip, though reduced, was relatively persistent.

Both discordance outcomes showed some fluctuations over the first few months, before a steady decline. This pattern is visually most notable in the graphs of normalized Z-scores presented in Fig. 3 (Z-RSD) and Fig. 4 (Z-RMSE), respectively. Within the first month (T1–2), ERT had a greater apparent impact on the severity of the facial discordance, compared to the proportion of discordance. Z-RSD and Z-RMSE scores both increased rapidly after BMT conditioning (T3–5) and peaked at around 4 months after BMT (T5, T6), during the BMT/ERT/cyclosporin phase. Progressive lessening of facial severity and discordance becomes apparent after 4–5 months (T5, T6), which corresponded to the BMT/ERT/cyclosporin treatment phase. This reduction continued until the final assessment time point. Overall, both Z-scores showed a decreasing trend, which was more pronounced in the Z-RMSE scores; Z-RMSE displayed larger scores (>3 SD higher than the mean) at its peak, in comparison to the Z-RSD scores that mostly fell within the normative range (max point: 1.5 SD from the mean).

Fig. 1 Method workflow used for the MPS I treatment-monitoring process

Discussion

For the first time we objectively, noninvasively, and dynamically assessed the changing 3-dimensional facial dysmorphology in a child undergoing disorder-specific treatment for a systemic metabolic disorder, namely, MPS I. This deeply precise dysmorphometric assessment demonstrated a reduction in facial dysmorphology which was in accordance with expectations from clinical experience and with that based on an exploratory cross-sectional study, demonstrating the correlation of the severity of facial and clinical MPS I phenotype (Kung et al. 2012). Additionally, it builds upon previous 3D facial analysis-based treatment monitoring of a localized facial pathology (Baynam et al. 2013) and of a dysmorphic non-metabolic disorder (de Souza et al. 2013). This study extends the findings of the aforementioned investigations, and they support that deep facial phenotyping may have applications for the development of treatment response biomarkers.

The greatest differences were found in the severity of facial discordance (root mean square error, RMSE). Additionally, overall, the longitudinal pattern of changes in the proportion of facial discordance (relative significant discordance; RSD) was in accordance with the pattern of the RMSE. This supports that approaches that provide multiple facial outcome measures may act to (1) corroborate each other, (2) provide complementary approaches to summarize variations in facial form, and (3) provide contrasting windows through which to consider the implications of found facial variation.

As treatments, including ERT and BMT, may ameliorate, but not reverse or prevent all MPS I manifestations (Wynn 2011), the expectation is of an initial period of improvement followed by a period of stabilization. Therefore, it is possible that there may be some persistence of residual facial dysmorphology and/or its partial recurrence. Our study's findings are in accordance with anticipated improvement/stabilization; however, it is possible that there may be persisting residual/recurrent facial dysmorphology over time.

The fluctuations in the global discordance scores detected during the first few months of therapy are notable. There was an initial reduction in facial severity during the initiation of ERT. Subsequently there was an overall ballooning facial effect, which may be related to transient subcutaneous facial fluctuations that were the result of the effects of conditioning and events in the initial post-BMT period. Should the documented variations be correlating with acute treatment events, this might then indicate our approach could be both a sensitive and rapid means of assessing treatment response or failure for established therapies, in drug development and clinical trials.

It is important to note that the measures (RSD, RMSE) used in this study are, by definition, relative scores. The purpose of the NE is to remove confounding factors like within-population variances (e.g., age, gender, ethnicity, and body mass index), which provides a more individualized assessment. It is the harmonic regions of the patient facial scan that drives the construction of the corresponding NE facial scan. Longitudinal changes, such as ageing, alter the patient's facial configuration at each time point, which

Fig. 2 MPS I longitudinal treatment monitoring over eight time points

Fig. 3 Progression of normalized facial discordance proportions along the MPS I treatment course. Z-relative significant discordance (Z-RSD; %; red points) scores were computed to enable standardization against the normative reference population

Fig. 4 Progression of normalized facial severity scores along the MPS I treatment course. Z-root mean squared error (Z-RMSE; mm; purple points) scores were computed to enable standardization against the normative reference range

then alters the harmonic interrelationships within that patient scan. This, in turn, changes the NE with each new time point. Hence, in this study we see eight similar yet different NE facial scans of the same patient. Therefore, direct comparisons of normal equivalents between time points require consideration within this context. For instance, the fullness of the labial region, on a background of the typical fuller face of infancy, might be more “harmonic” and therefore less discordant when compared to labial fullness in an older child. Interestingly, this labial fullness is seen in the 14-year-old MPS IH individual reported by Kung et al. (2012), where lip prominence was detected as a discordant feature 13 years after BMT. This unmasking of present, but challenging to discern without objective support, and age-dependent dysmorphology could partly explain challenges in unaided clinical facial assessment. However, they may be identified with objective and precise approaches such as the one described herein.

This study was limited to the assessment of one individual, and other treatment endpoints were not available for comparison. Further studies will be required to investigate additional individuals and importantly to relate their facial phenotype to existing outcome measures. Given the individual rarity of these conditions, this will require multi-institutional assessments, e.g., urinary glycosaminoglycans and measures of respiratory function and endurance. Fortunately, the wide geographic dispersion of currently available robust and precise imaging equipment can facilitate this in the immediate term. Also, rapid advances in the development of increasingly portable and less costly 3D imaging equipment will make this increasingly feasible in the intermediate term including for point of care or ultimately in-home image capture.

This approach could be expanded to other metabolic and non-metabolic disorders with known (or as yet unappreciated) facial phenotypes with current, and emerging, disorder-specific therapies.

Conclusion

This longitudinal study demonstrated objective and deeply precise changes in facial morphology in a child with treated MPS I. If corroborated by further studies, including correlation with other objective treatment outcomes, this supports the use of dysmorphometric 3D facial analysis as a noninvasive and relatively inexpensive treatment biomarker to rapidly assess therapeutic response in this condition and possibly other disorders with facial dysmorphology. Additionally, our longitudinal findings in a young child,

particularly when coupled with previous cross-sectional studies (Kung et al. 2012), suggest that this approach could aid early clinical classification.

Acknowledgments We are indebted to the child with MPS I and his family for participation in this study and permission for image use. We would like to thank Genzyme Australia for providing an unrestricted educational grant. The Princess Margaret Hospital Foundation in Perth, Western Australia, also provided support for this study. RD-Connect, the associated National Health and Medical Research Council of Australia, and the Raine Clinician Research Fellowship supported GB’s contribution.

Synopsis

Longitudinal quantification of objective and deeply precise changes in MPS I facial morphology during treatment demonstrates the potential of dysmorphometric 3D facial analysis as a noninvasive treatment response biomarker.

Compliance with Ethics Guidelines

Conflict of Interests

Stefanie Kung (SK), Mark Walters (MW), Peter Claes (PC), Peter LeSouef (PLS), Jack Goldblatt (JB), Andrew Martin (AM), Shanti Balasubramaniam (SB), and Gareth Baynam (GB) declare that they have no conflict of interests. Genzyme provided an unrestricted educational grant and had no role in the interpretation/analysis of the data or the decision to submit the manuscript for publication.

Informed Consent

All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2000 (5). Informed consent was obtained from all participants included in the study.

Author Contributions

SK drafted the initial manuscript with subsequent revisions and input from **MW**, **GB**, **JG**, and **PC**. **GB** and **JG** conceived the study. **SK** acquired the longitudinal facial data and implemented the facial assessments. **PC** performed the facial mapping of the data and generated normative reference discordancy statistics. All authors were involved in the interpretation of the results. **SB**, **AM**, **JG**, and **PLS**

provided valuable clinical insight into the patient's treatment regime, and the investigated syndrome.

Guarantor

GB accepts full responsibility for this work and conduct of this study as guarantor for this article. He has had access to the research data and controls the decision to publish.

References

- Aldridge K, Boyadjiev SA, Capone GT, DeLeon VB, Richtsmeier JT (2005) Precision and error of three-dimensional phenotypic measures acquired from 3dMD photogrammetric images. *Am J Med Genet A* 138A(3):247–253
- Baynam GS, Walters M, Dawkins H, Bellgard M, Halbert AR, Claes P (2013) Objective monitoring of mTOR inhibitor therapy by three-dimensional facial analysis. *Twin Res Hum Genet* 16(4):840–844
- Church H, Tylee K, Cooper A et al (2007) Biochemical monitoring after haemopoietic stem cell transplant for Hurler syndrome (MPSIH): implications for functional outcome after transplant in metabolic disease. *Bone Marrow Transplant* 39(4):207–210
- Claes P (2007) A robust statistical surface registration framework using implicit function representations: application in craniofacial reconstruction. KU Leuven, Leuven
- Claes P, Daniel K, Walters M, Clement JG, Vandermeulen D, Suetens P (2012a) Dymorphometrics: the modelling of morphological anomalies. *Theor Biol Med Model* 9(5):1–39
- Claes P, Walters M, Clement JG (2012b) Improved facial outcome assessment using a 3D anthropometric mask. *Int J Oral Maxillofac Surg* 41:324–330
- Claes P, Walters M, Gillett D, Vandermeulen D, Clement JG, Suetens P (2013) The normal-equivalent: a patient-specific assessment of facial harmony. *Int J Oral Maxillofac Surg* 42(9):1150–1158
- Clarke LA, Heppner J (2011) Mucopolysaccharidosis type I. University of Washington, Seattle
- de Souza MA, McAllister C, Suttie M et al (2013) Growth hormone, gender and face shape in prader-willi syndrome. *Am J Med Genet A* 161(10):2453–2463
- Hammond P, Suttie M (2012) Large-scale objective phenotyping of 3D facial morphology. *Hum Mutat* 33(5):817–825
- Kung S, Walters M, Claes P, Goldblatt J, Le Souef P, Baynam G (2012) A dymorphometric analysis to investigate facial phenotypic signatures as a foundation for non-invasive monitoring of lysosomal storage disorders. Springer, Berlin/Heidelberg, pp 1–9
- Muhlstein A, Gelperina S, Kreuter J (2013) Development of nanoparticle-bound arylsulfatase B for enzyme replacement therapy of mucopolysaccharidosis VI. *Pharmazie* 68(7):549–554
- van Gelder CM, Vollebregt AA, Plug I, van der Ploeg AT, Reuser AJ (2012) Treatment options for lysosomal storage disorders: developing insights. *Expert Opin Pharmacother* 13(16):2281–2299
- Wynn R (2011) Stem cell transplantation in inherited metabolic disorders. *Hematology* 2011:285–291