J Inherit Metab Dis (2011) 34:1053–1059
DOI 10.1007/s10545-011-9347-z

ORIGINAL ARTICLE

Four-year follow-up of chronic neuronopathic Gaucher disease in Europeans using a modified severity scoring tool

Elin Haf Davies · Eugen Mengel · Anna Tylki-Szymanska · G. Kleinotiene · Joerg Reinke · Ashok Vellodi

Received: 30 November 2010 / Revised: 27 April 2011 / Accepted: 3 May 2011 / Published online: 28 May 2011

Abstract In 2007, the European Task Force for neuronopathic Gaucher disease (NGD) published a review of 55 patients across four countries. Although some observations were possible, analysis was difficult due to the absence of a systematic way of assessing patients. In response to this, a Severity Scoring Tool (SST) was devised to offer a systematic means of assessing the neurological presentation seen. The SST has been modified (mSST) and is a valid tool for monitoring neurological progression. This review describes disease status and progression of neurological manifestations in a cohort of 39 chronic NGD patients across three European countries over a period of 4 years, using the mSST.

Communicated by: Gregory M. Pastores

Competing interest: None declared

E. H. Davies (✉)
Institute of Child Health,
London, UK
e-mail: E.Davies@ich.ucl.ac.uk

E. Mengel · J. Reinke
Villa Metabolica Childrens Hospital, Universitatsmedizin, Mainz,
Langenbeckstr, Germany

A. Tylki-Szymanska
The Children's Memorial Health Institute,
Warsaw, Poland

G. Kleinotiene
Vilnius University Children's Hospital,
Vilnius, Lithuania

A. Vellodi
Great Ormond St Hospital for Children NHS Trust,
London, UK

Introduction

The original EU Task Force guidelines for neuronopathic Gaucher disease (NGD) management were published in 2001 (Vellodi et al. 2001). In 2007, in an attempt to evaluate the effect of high-dose enzyme replacement therapy (ERT) on the neurological outcome of NGD patients, a review of patients in four European countries was conducted (Davies et al. 2007a). Despite a systematic, retrospective collection of data, analysis was difficult due to the heterogeneity of patients and the absence of systematic assessment. Some observations were possible, however, and as a result, revised recommendations were issued by the task force (Vellodi et al. 2009). In parallel, a Severity Scoring Tool (SST) for NGD was developed (Davies et al. 2007b) with the aim of providing a systematic means of assessing patients that was user friendly and without cultural or economic constraints. The SST is an 11-domain tool accounting for the neurological manifestations observed in NGD and yielding a score ranging from 0 to 33 according to disease severity. The SST was incorporated in the revised recommendations for monitoring NGD (Vellodi et al. 2009). Following further formal consultation with seven international experts in the form of two separate Web-Ex conferences, the SST was further validated and modified and renamed the modified Severity Scoring Tool (mSST) (Table 1). Modification resulted in adjustment of scores attributed to each neurological domain based on the impact on disease severity. Whereas previously all domains had a maximum score of 3, the epilepsy and pyramidal domains now scored highest (5), whereas horizontal gaze palsy (HGP) scored the least (1). A 12th domain was also added to the tool—age at first seizure—which was incorporated based on expert comments and feedback to distinguish between patients with severe neurological

Table 1 Modified Severity Scoring Tool (mSST)

Gaze palsy	Normal (although not likely in diagnosis)	0
	Horizontal saccades absent, vertical saccades present	0.5
	Horizontal saccades and vertical saccades absent	1
Ophthalmology	Normal	0
	Cranial nerve palsy (previously corrected or not)	1
	Cranial nerve palsy (reappearing despite surgical correction)	2
Epilepsy	No seizures.	0
	Seizures not requiring anticonvulsants	3
	Seizures controlled with anticonvulsants.	4
	Seizures requiring combination therapy or resistant to anticonvulsants	5
Age at first seizure	Younger than 5 years	3
	5–10 years	2
	10–15 years	1
	16 years or over, or seizure free	0
Development/ Cognitive ability	Normal	0
	Mildly impaired (IQ less than 85 or equivalent)	1
	Moderate (IQ 50–57 or equivalent)	2
	Severe (more than half their chronological age)	3
Ataxia of gait	Normal, apparent only on tandem walking	0
	Ataxia on straight gait, able to walk without assistance	1
	Able to walk only with assistance	2
	Unable to walk	3
Cerebellar tremor	No intention tremor	0
	Intention tremor not affecting function	0.5
	Intention tremor with marked impact on function	2
Pyramidal	Normal tone with increased reflexes	0
	Mildly to moderately increased tone and reflexes	2
	Increased tone and reflexes with clonus, whether unsustained or sustained	3
	Severe spasticity with inability to walk	5
Extrapyramidal	Normal	0
	Variable tone and posturing not impairing function, with or without therapy.	1
	Variable tone and posturing impairing function, despite therapy	2
	Significant rigidity with no/minimal benefit from therapy	3
Swallowing difficulties/ Oral bulbar function	Normal	0
	Mild dysphagia (excess drooling)	1
	Moderate dysphagia (risk of aspiration, modification to diet required)	2
	Severe dysphagia (requiring non-oral feeding)	3
Speech	Normal (and those too young yet to speak)	0
	Mild to moderate dysarthia impairing intelligibility to unfamiliar listener	1
	Severe dysarthia with most speech unintelligible to familiar and unfamiliar listener	2
	Anarthria	3
Spinal alignement	Normal	0
	Mild kyphosis–but flexible and not requiring bracing	1
	Moderate kyphosis–partially corrected by bracing	2
	Severe kyphosis–not corrected by bracing or requiring surgery	3
Total		36

disease at a young age, as this was felt to be a clinical hallmark of early severe neurological presentation.

Definition

Consistent with the original review, NGD was defined as the presence of neurological involvement in patients with biochemically proven Gaucher disease for which there was no explanation other than Gaucher disease. For the purposes of this paper, acute/type II was defined by the onset in infancy of progressive bulbar involvement as stated in the original European Task Force guidelines (Vellodi et al. 2001). Acute NGD (type II) patients were not studied. All other NGD patients were assumed to have chronic NGD, as the duration of follow-up was 4 years, and that patients with acute NGD rarely live beyond 4 years of age.

Data collection

Basic demographics: sex, ethnicity, genotype, and spleen status were available from the original review. A mixture of polymerase chain reaction (PCR), sequencing, and Southern blots were used to determine genotype. Follow-up data from patients assessed in the original review were collated from three European centers: Mainz (Germany), Warsaw (Poland), and London (United Kingdom). Data collected at follow-up were:

- Age at assessment
- Seizure history
- mSST scores
- Neuropsychometric scores
- Current ERT dose
- Recent chitotriosidase levels

Visceral manifestations other than chitotriosidase were not captured as part of this review. All data were collated and analyzed by E. H. Davies. Each expert assessed each patient in their respective country.

Statistical methods

Descriptive tests (mean and median) were used to describe demographic characteristics, mSST, neuropsychometrics, ERT, and chitotriosidase. All SST scores were converted to mSST scores. Nonparametric tests were used to analyze change to account for the small sample sizes in each genotype group. A multiple regression analysis was performed to predict mSST score at follow-up, accounting for baseline mSST, full-scale IQ, and genotype.

Results

Demographic data

Data are summarised in Table 2. There were 39 patients in all, and all were chronic (cNGD) in presentation; no acute/type II patients were included in this review, as per previous definition. Distribution was as follows: Poland 18 (included two Lithuanian patients), Germany 10, UK 11. The original 2007 review included Swedish patients. However, for logistical reasons, it was not possible to include patients from Sweden in this subsequent review. The Polish and British cohorts had the oldest and youngest average ages, respectively.

Consistent with the original review, the majority of patients were homozygous for the L444P mutation (n=27, 69.2%). This overall incidence of 69.2% is similar to the 72% reported in the International Collaborative Gaucher Group Neurological Outcomes Subregistry (Tylki-Szymańska et al. 2010). However, there was a much lower percentage noted in Germany (50%) than in Poland and the UK (77.7% and 72.7%, receptively). For analysis purposes, the remaining genotypes were grouped to those with one L444P allele and one other; F213I/L444P, L444P/G202R, L444P/E326K, and L444P/E233D. Heterozygotes for the D409H and L444P alleles and all the remaining genotypes were classified as other: L279P/G243V, R433S/R433S, D409H/G202R,

Table 2 Demographic data of European chronic neuronopathic Gaucher disease (cNGD) cohort assessed

	Poland	Germany	UK	Total
Number	18	10	11	39
Mean age (years) at baseline	19.2 (SD±11.1)	13.8 (SD±8.8)	9.7 (SD±4.35)	15.1 (SD±9.8)
Mean age (years) at follow-up	23.2 (SD±11.2)	17.9 (SD±8.8)	13.1 (SD±4.6)	18.5 (SD±9.9)
Median age (years) at baseline	19.6	11.9	9.3	11.9
Median age (years) at follow-up	15.7	16.0	13.3	16.0
L444P/L444P	14	5	8	27 (69.2%)
L444P/other allele	1	1	2	4 (10.2%)
L444P/D409H	2	2	0	4 (10.2%)
All other Genotypes	1	2	1	4 (10.2%)
Total/partial splenectomy	6	2	0	8 (20.5%)

SD standard deviation

Table 3 Enzyme replacement therapy and chitotriosidase data

	Poland	Germany	UK	Complete cohort	Significant
Median baseline ERT (IU/kg per 2 weeks) (*n* 34)	30.0 (26–59)	87.2 (64–117)	117 (101–173)	70.4 (30–107)	0.000*
Median follow-up ERT (IU/kg/2 weeks) (*n* 34)	26.5 (16–44)	56.5 (44–69)	114 (102–009)	51.0 (24.8-84)	
Median baseline chitotriosidase (nmol/h/ml) (*n* 33)	2260 (710–6,128)	1185 (423–3,575)	787 (520–2,280)	1544 (542–3,562)	0.020*
Median Follow-up chitotriosidase (nmol/h/ml) (*n* 33)	547 (330–2,185)	1145 (265–3,429)	557 (388–1,700)	625 (340–2,278)	

* Statistically significant

1599AG/1603T. The young German cohort had a high proportion of compound heterozygotes, which might indicate the more severe disease compared with the older, primarily L444P homozygote, Polish cohort. No patients underwent a total splenectomy during the follow-up period.

Treatment

Chitotriosidase and ERT data are summarised in Table 3. All patients were receiving ERT. However, doses were significantly affected by the current worldwide shortage, but which might also reflect the revised recommendations regarding therapy (Vellodi et al. 2009).

Despite the reduced dose, the median chitotriosidase was lower, reaching statistical significance, possibly indicating that time on ERT rather than ERT dose impacts chitotriosidase levels, although the impact of reduced dose may yet to be seen. This is evident for the Polish cohort, for example, where the largest reductions in chitotriosidase levels were seen despite patients being on the lowest dose. Of note, these patients were also milder in terms of presenting neurological manifestations, as scored with the mSST, at both baseline and follow-up, illustrating that there is no link between seen chitoriosidase and neurological profile.

Seizures

The number of patients with seizures increased from four to eight during the 4-year period. At follow-up, this equates to 20% of the cohort, which is not too different to the 16% reported in the NGD registry publication (Tylki-Szymańska et al. 2010). Genotypes of those presenting with seizures at follow-up were: L279P/G243V, F213I/L444P, D409H/G202R—the remaining five being L444P homozygote. None of the D409H/L444P cohort presented or developed seizures. At baseline, only two of those with seizures were L444P homozygote, which was just 7%; at follow-up, this is increased to 18%. Only one of the eight patients experienced myoclonic seizures, whereas the others experienced tonic clonic and/or absences. The number of patients with seizures requiring combination therapy or resistant to antiepileptic drugs (AED), as defined in the

Table 4 Modified Severity Scoring Tool scores at baseline and follow-up (Median and Q25/Q75)

	Baseline (Q25/Q75)	Follow-up (Q25/Q75)	Significant *p*
Poland (*n*=18)	3.5 (1.9–5.6)	5.0 (2–6.5)	0.152
Germany (*n*=10)	5.5 (2.3–9.8)	8.5 (2.6–16.1)	0.043*
UK (*n*=11)	6.0 (2–12)	6.0 (1.5–12.5)	1.77
Cumulative cohort (*n*=39)	4.0 (2–6.5)	6.0 (2–10)	0.007*
L444P/L444P (*n*=27)	4.0 (2–6.5)	6.0 (2–7)	0.032*
L444P one allele (*n*=4)	9.5 (2.6–13.4)	12 (3.5–19.4)	0.465
L444P and D409H (*n*=4)	2.8 (2.1–3.4)	3.0 (1.1–4.1)	1.000
Other genotypes (*n*=4)	8.0 (3–11.1)	13.3 (3.6–23.6)	0.109
Age at follow-up=< 18 years (*n*=24)	3.5 (2–6.5)	6.3 (1.6–9.8)	0.018*
Age at follow-up => 18 years (*n*=15)	5.5 (2–8.5)	5.0 (3–12)	0.178

*Statistically significant

mSST epilepsy domain, increased from one to five, indicating an increased seizure severity This is in keeping with the known progressive nature of the seizures seen in NGD .

Neuropsychometric assessments

The Full-scale Intelligence Quotient (FSIQ) was captured, as this was identified as the only consistent assessment performed in the countries during the original review. The tests used varied depending on patient age but included the Wechsler Intelligence Scale for Children (WISC-III) and Wechsler Adult Intelligence Scale–III (WAISI-III) or its predecessor. Follow-up FSIQ was only captured in 14 patients. Unfortunately, verbal and performance IQ were not available separately in all patients, so it was not possible to say whether there was any discrepancy. Of the 14 patients for whom FSIQ was available, 12 were L444P homozygote, one was R433S/R433S, and one was D409H/L444P. Indeed, both these patients scored a high FSIQ of 97 and 96, respectively, which is in the average IQ bracket. At follow-up, the R433S/R433S patient had an improved FSIQ score of 108, whereas the D409H/L444P maintained the same score. The mean baseline FSIQ for the L444P homozygote cohort was 80 (±15.3), which is on the borderline of the low/ borderline learning difficulties and low average brackets. This increased to 82.75 (±20.3) at follow-up. However, this increase was neither statistically nor clinically significant, as an FSIQ of 83 is still within the low average bracket of function. Collectively as a group of 14, baseline FSIQ actually increased from 82.36 (±15.2) to 85.5 (±20.1) by follow-up. Again, this was neither statistically nor clinically significant.

Modified Severity Scoring Tool (mSST) scores

The median baseline mSST score (calculated by converting the original SST score) was 4.0 (2–6.5). There is a difference in score across the genotypes as grouped: D409H and L444P allele group scores 2.75 (2.1-3.4); L444P homozygote group scores 4.0 (2–6.5); other genotypes scores 8.0 (3–11). The L444P/other allele group was the highest scoring genotype group at 9.5 (2.6–13.4) (Table 4 and Fig. 1).

Median follow-up mSST score for the cohort collectively increased to 6.0 (2–10). Again, this differed greatly across the various genotypes: 3.0 (1.1–4.1) in those with a D409H and L444P allele; 6.0 (2–7) in the L444P homozygote group; 12.0 (3.5–19.4) in those with only one L444P allele; 13.3 (3.6–23.6) in the other genotypes cohort. The large distribution in the last two groups is a reflection of the heterogeneity seen in these patients (Table 4 and Fig. 2).

Using a Wilcoxon test to account for the small sample size and data distribution, the change in mSST score across the entire cohort was statistically significant (p 0.007, Table 4). Again, the increase in mSST score varied across genotypes: two in the L444P homozygote group; 2.5 for those with one L444P allele; 5.3 for the other genotypes cohort. There was a minimal change in the D409H/L444P cohort of 0.2. Clinically, this suggests that the D409H/L444P is associated with a mild phenotype. However, care needs to be taken in making this assumption, as the number is small. When testing the significance in score change according to genotype, only the L444P/L444P cohort was statistically significant (p. 0.032). The group categorised as other genotypes demonstrated the greatest increase in mSST score; however, this increase was not significant. This was probably due to the small sample size and considerable heterogeneity.

The German cohort demonstrated the greatest increase in mSST score. This suggests that the two things that set this cohort apart—i.e., the high proportion of compound heterozygotes and the much younger average age—are responsible. In contrast, the Polish cohort, characterised by a high proportion of L444P homozygotes and an older average age, yielded narrower and more consistent quartiles. We wanted to determine whether Polish L444P

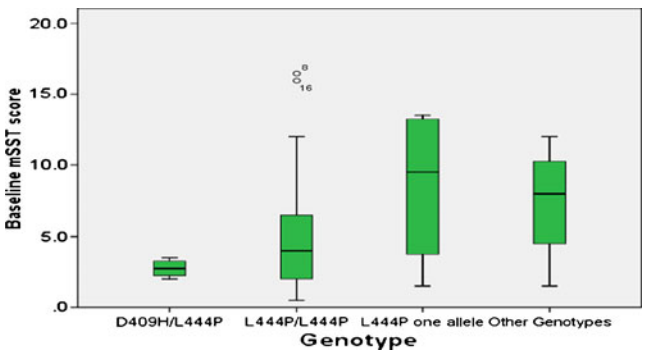

Fig. 1 Baseline Modified Severity Scoring Tool (mSST) scores across genotypes (n 39)

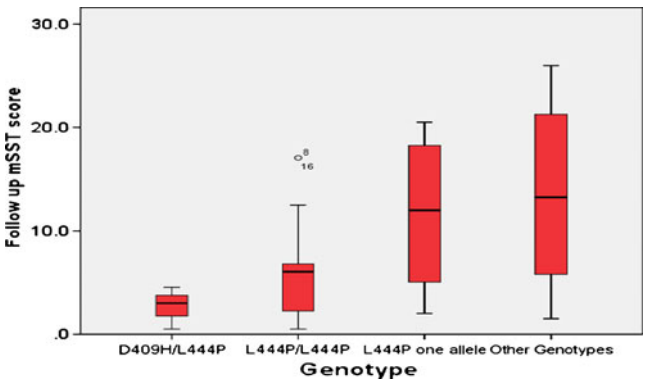

Fig. 2 Follow-up Modified Severity Scoring Tool (mSST) score across genotypes (n 39)

Table 5 Modified Severity Scoring Tool scores for the ten patients who demonstrated improvement in score, according to substrate reduction therapy

Substrate reduction therapy	Baseline mSST score Median (25–75)	Follow-up mSST score Median (25–75)
No ($n=6$)	4.75 (3–6.9)	3.25 (0.5–6.1)
Yes ($n=4$)	4.25 (2–11.4)	3.75 (1.5–7.5)

patients (14 of 18) were atypical for this genotype. It can be seen from Table 4 that baseline and follow-up mSST data for the Polish cohort are, in fact, comparable with those for the entire L444P homozygote cohort. This indicates that all L444P homozygote patients, at least in this European cohort, have a comparable neurological severity. The UK cohort remained stable during this review period, with both median mSST scores and quartiles remaining constant. As previously identified, the UK cohort was genotypically similar to the Polish cohort in terms of percentage of L444P homozygote patients, but with a much younger age range. This high percentage of L444P homozygote patients may explain this stability compared with the heterogeneous cohort found in Germany,.

When exploring data according to age brackets (<18 years vs >18 years at the time of follow-up assessment), those <18 years demonstrated a clear increase in mSST score, which was statistically significant (p 0.018). By contrast, patients >18 years appeared to be stable. This may reflect the fact that the older patients generally had a milder presentation of disease, hence the longer survival.

Despite the overall increase in mSST score across the cohort, ten patients reported an improved score. Five patients only improved by 0.5, which was based on improvement in cerebellar tremor (intention). Four of these five patients were on substrate reduction therapy (SRT) at the time of baseline assessment (Schiffmann et al. 2008). This may explain the small improvement in score, as the known side effect of SRT is intention tremor, which resolved on stopping therapy. However, improvement in mSST scores at follow-up was noted in two more patient not treated with SRT (Table 5), and given the small sample size, no conclusions can be drawn.

Individual domain analysis

An analysis of each domain was performed to identify which one showed the greatest change. The percentage of patients presenting with gaze palsy remained constant, as would be expected, and for pyramidal. Involvement of cognitive ability, ataxia/gait, speech, and extrapyramidal increased at a similar rate of 2.5%. Epilepsy and kyphosis increased by 5%, whereas cerebellar signs/ataxia involvement was nearly 13% higher. Two domains, swallowing and ophthalmology, saw a 2.5% reduction in involvement (Fig. 3). These findings indicate the importance of using a panel of neurological domains rather than just one as a marker of neurological status. Percentage of patients (n=39) presenting with each individual domain at baseline and follow-up are shown in Fig. 3.

Value of mSST in predicting disease course

The value of the mSST in measuring neurological change was then explored. The independent variables considered by the panel most likely to influence the score were:

- Baseline mSST
- Genotypes (categorised into dummy variables)
- Spleen status (categorised into yes/no)
- Age at first assessment
- FSIQ at baseline

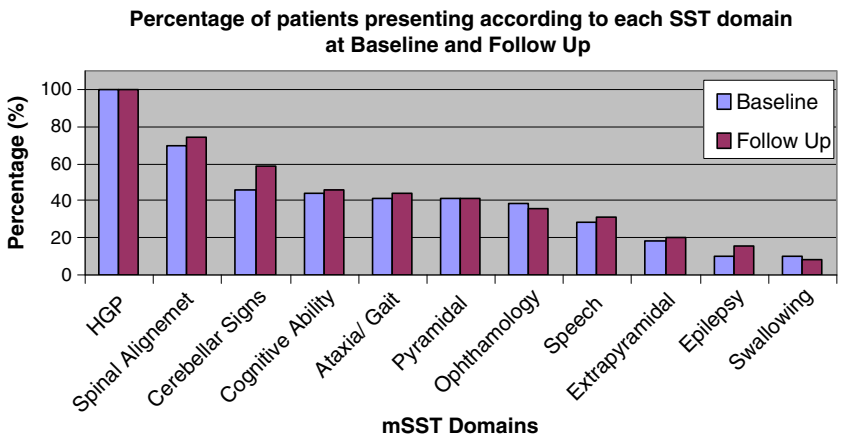

Fig. 3 Percentage of patients (n=39) presenting with each individual domain at baseline and follow-up

With these variables, R^2 of 0.727 was achieved. Using the multiple regression equation to predict the follow-up mSST score for any given patient, the following model can be proposed, where L444P/L444P is baseline:

$$\text{Follow-Up mSST score}' = 5.485 + -0.151 \times (Age) + -0.037 \times (FSIQ) + 0.999 \times (Baseline\ mSST) + 7.712 \times (D409H/L444P) + -0.492 \times (Other\ Genotypes) + 2.077 \times (Splenectomized).$$

(Standard error of the slope ±3.458; 95% confidence interval for the slopes −1.9 to 12.9). The mSST score at baseline makes the greatest contribution to predicting follow-up mSST score; a change of one standard deviation on that variable produces a change of 0.865 on follow-up mSST score. In fact, the baseline mSST score is the only significant predicting variable.

Discussion

The primary objective of this review was to determine whether the mSST was sensitive enough to capture disease progression and contribute further to our understanding of the natural history of cNGD in the ERT era. This is the largest cohort of cNGD patients ever assessed prospectively and systematically over a period of 4 years, which is long enough to capture the progressive nature of the disease. There were limitations to the study:

- The relatively small number of patients
- The considerable heterogeneity; in particular, the younger German cohort had a high proportion of compound heterozygotes (which might indicate more severe disease) compared with the primarily L444P homozygote, older, Polish cohort

Nevertheless, we believe that the mSST captures disease progression. It also broadly distinguishes between genotypes that have a phenotypic correlation. The relatively small increase in score cumulatively is reflective of the slow nature of disease progression in the majority of the cohort. Progression in heterozygote patients is more heterogeneous but is generally associated with increased disease severity and greater progression.

Conclusion

The mSST is sensitive to capturing change, is user friendly, and has no cultural or economic constraints. Its wide-spread use in the management of NGD patients, particularly beyond Europe, could serve as a means of generating a greater understanding of disease progression and ultimately serve as a basis for the design of future clinical trials.

Acknowledgement This work was made possible by grants from the Gaucher Association UK and Genzyme. Thank you also to the international experts who took part in the discussion to modify the SST: Amal El Beshlawy, Egypt; B. Bembi, Italy; Edward Giannini, USA; Thong Meow Keong, Malaysia; Ana Maria Martins, Brazil; A Tylki-Szymanska, Poland; Ashok Vellodi, UK.

References

Davies E, Erikson A, Collin-Histed T, Mengel E, Tylki-Szymanska A, Vellodi A (2007a) Outcome of type III Gaucher disease on enzyme replacement therapy: review of 55 cases. J Inherit Metab Dis 30(6):935–42

Davies E, Surtees R, DeVile C, Schoon I, Vellodi A (2007b) A severity scoring tool to assess the neurological features of neuronopathic Gaucher disease. J Inherit Metab Dis 30(5):768–82. 5, Epub 2007 Sep 16

Schiffmann R, Fitzgibbon EJ, Harris C, DeVile C, Davies EH, Abel L, van Schaik I, Timmons M, Ries M, Vellodi A (2008) A randomized controlled trial of miglustat in Gaucher disease type 3. Ann Neurol 64(5):514–22

Tylki-Szymańska A, Vellodi A, El-Beshlawy A, Cole JA, Kolodny E (2010) Neuronopathic Gaucher disease: demographic and clinical features of 131 patients enrolled in the International Collaborative Gaucher Group Neurological Outcomes Subregistry. J Inherit Metab Dis 33(4):339–46

Vellodi A, Bembi B, de Villemeur TB, Collin-Histed T, Erikson A, Mengel E, Rolfs A, Tylki-Szymanska A (2001) Neuronopathic Gaucher Disease Task Force of the European Working Group on Gaucher Disease. 2001 Management of neuronopathic Gaucher disease: a European consensus. J Inherit Metab Dis 24(3):319–27

Vellodi A, Tylki-Szymanska A, Davies EH, Kolodny E, Bembi B, Collin-Histed T, Mengel E, Erikson A, Schiffmann R (2009) Management of neuronopathic Gaucher disease: revised recommendations. J Inherit Metab Dis 32(5):660–4, Aug 5