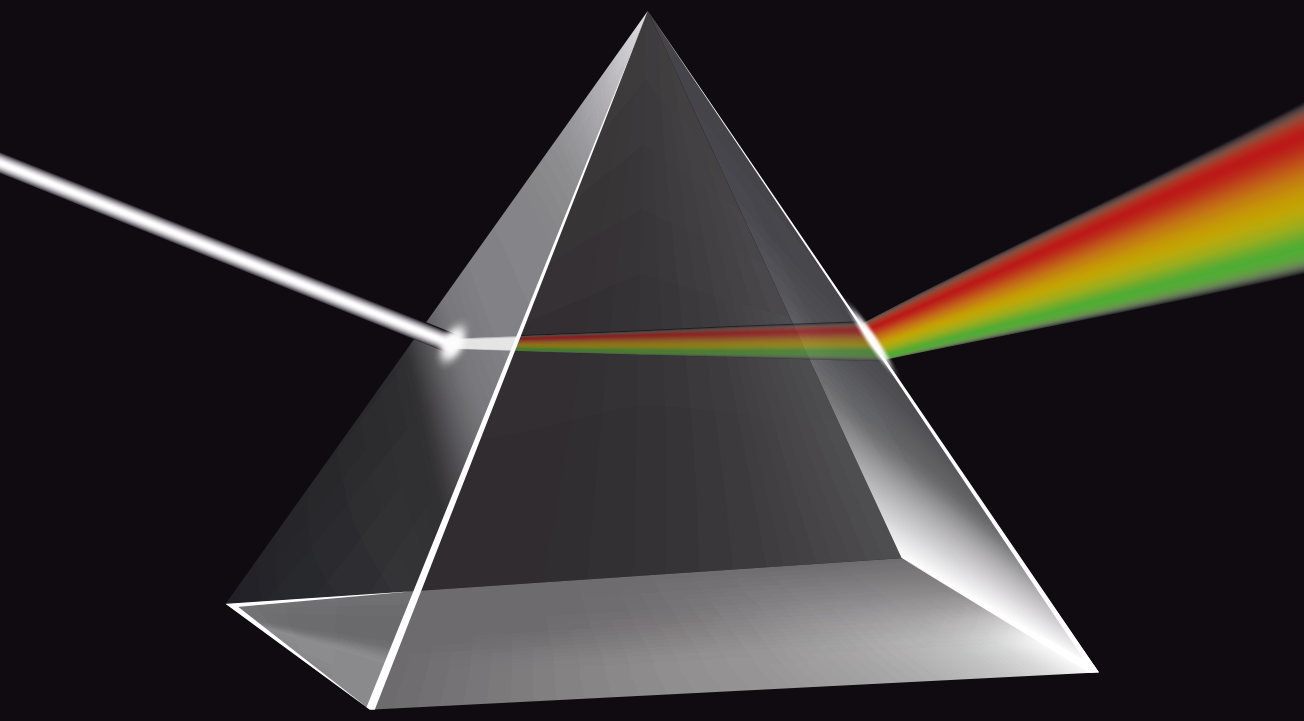
Classical Galactosemia

Elucidating the spectrum of clinical outcome

Mendy Welsink-Karssies



Classical Galactosemia

Elucidating the spectrum of clinical outcome

Mendy Welsink-Karssies

Classical Galactosemia - Elucidating the spectrum of clinical outcome
Author: Mendy Welsink-Karssies
ISBN: 978-94-6423-026-0

Financial support for printing of this thesis was kindly provided by Galactosemie Vereniging Nederland (GVN), Castor and Nutricia.

Cover image and design by Dennis Hendriks || www.proefschriftmaken.nl
Layout and printed by ProefschriftMaken || www.proefschriftmaken.nl

Classical Galactosemia

Elucidating the spectrum of clinical outcome

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor

aan de Universiteit van Amsterdam

op gezag van de Rector Magnificus

prof. dr. ir. K.I.J. Maex

ten overstaan van een door het College voor Promoties ingestelde commissie,

in het openbaar te verdedigen in de Agnietenkapel

op vrijdag 22 januari 2021, te 16.00 uur

door

Mendy Marcella Karssies

geboren te Amsterdam

PROMOTIECOMMISSIE

Promotor:	prof. dr. F.A. Wijburg	AMC-UvA
Copromotor:	dr. A.M. Bosch	AMC-UvA
Overige leden:	prof. dr. J.B. van Goudoever	AMC-UvA
	prof. dr. R.H.L. Houtkooper	AMC-UvA
	prof. dr. J. Oosterlaan	Vrije Universiteit Amsterdam
	prof. dr. M.S. van der Knaap	Vrije Universiteit Amsterdam
	dr. M. Langeveld	AMC-UvA
	dr. E. Murphy	National Hospital for Neurology and Neurosurgery

Faculteit der Geneeskunde

TABLE OF CONTENTS

Chapter 1

General introduction

GENERAL INTRODUCTION

Classical Galactosemia: introduction and history

Galactose metabolism

Classical galactosemia is an autosomal recessive inborn error of galactose metabolism in which the ability to metabolize galactose is severely hampered. Galactose is produced by the conversion of lactose into glucose and galactose by the enzyme lactase (EC 3.2.1.108). In infancy, lactose is the primary dietary carbohydrate as it is the main carbohydrate in (breast) milk and dairy products[1]. Galactose is normally rapidly metabolized into glucose for energy by the Leloir pathway (**Figure 1**)[2], which comprises four different enzymes; (1) Galactose mutarotase (GALM, EC 5.1.3.3), (2) Galactose kinase (GALK, EC 2.7.1.6), (3) Galactose-1-phosphate uridylyltransferase (GALT, EC 2.7.7.12) and (4) UDP-galactose 4-epimerase (GALE, EC 5.1.3.2)[3]. Variations in genes that encode these enzymes can lead to enzyme deficiencies. Galactosemia, *i.e.* galactose in the blood, can be caused by a deficiency in one of the four different enzymes of this pathway.

A severe deficiency of the enzyme galactose-1-phosphate uridylyltransferase (GALT), caused by variations in the *GALT* gene (NM_000155.3) located on chromosome 9p13[4], leads to the disorder Classical Galactosemia (OMIM# 230400). Currently, 363 *GALT* variants have been reported[5,6]. The GALT enzyme converts galactose-1-phosphate (Gal1P) with uridine diphosphate-glucose (UDP-glc) to glucose-1-phosphate (G1P) and UDP-galactose (UDP-gal) (**Figure 1**)[2].

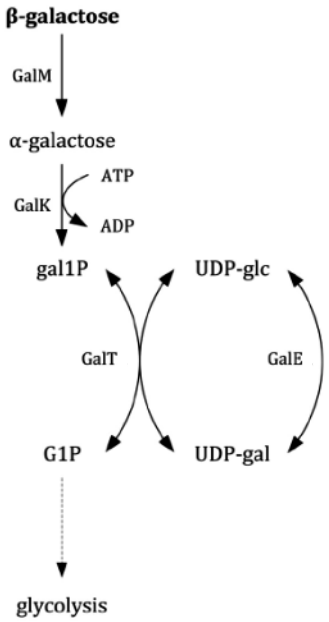


Figure 1. The Leloir pathway

History

The first patient with Classical Galactosemia (CG) was described in the case report "Sugar Excretion in Infancy" and reported on a breast-fed infant with failure to thrive, liver enlargement and 'galactosuria'[7]. The excretion of galactose in the urine disappeared after the removal of milk products from the diet. Hereafter, this disorder was treated by eliminating milk and dairy products from the diet, but it would take many years before the deficiency of the GALT enzyme activity was discovered in 1956[8]. Leloir and colleagues elucidated the galactose metabolic pathway in the early 1950's and Leloir won the Nobel prize (for Chemistry) in 1970 for his work[9]. In the early 1960's the first newborn screening (NBS) test for galactosemia was introduced in the United States[10]. Since then, CG has been included in the NBS of many countries[11,12].

Classical Galactosemia in the Netherlands

CG was included in the Dutch NBS in 2007. Before 2007, patients were diagnosed based on the clinical presentation (see below) or as a result of family screening in patients having an older sibling with CG. As CG is rare, a diagnosis in an index patient was usually made only after the exclusion of other more prevalent disease, leading to a diagnostic delay and often severe morbidity and significant mortality. Between 2007 and 2015, a total of 31 patients with CG were diagnosed out of 1.637.733 screened newborns, resulting in an incidence of 1:52.830[13]. CG is one of the more prevalent inborn errors of metabolism with currently 180 registered patients in the Dutch Diagnosis Registration Metabolic Diseases (DDRMD) since 1920[14]. There are two national recognized Galactosemia Expertise Centers in The Netherlands; the Amsterdam University Medical Center (UMC, location AMC) and the Maastricht University Medical Center (MUMC)[15]. Patients receive a standardized treatment and follow-up as recommended by the International guideline for the management of classical galactosemia[16].

Clinical presentation

In affected newborns, the ingestion of galactose from breastmilk or infant formula generally leads to severe illness in the first weeks of life. Patients may present with feeding difficulties, vomiting and jaundice followed by (life-threatening) illness affecting multiple organs such as liver failure with bleeding diathesis, renal tubular dysfunction, cataract, (*E. Coli* induced) sepsis and meningitis. If galactosemia is suspected in a newborn, a galactose-restricted diet (e.g. soy-based, casein hydrolysate or elemental formula) is started immediately, before conformation of the diagnosis by enzymatic or mutational analysis[16]. After the initiation of a galactose-restricted diet (life-threatening) symptoms resolve quickly. Screening for CG followed by an early initiation of galactose restriction is lifesaving in the newborn period and prevents critical illness in a (vast) majority of newborns[13,16,17]. Therefore, several countries have included CG in their NBS program[11,12].

Diagnosis and treatment

The diagnosis CG is confirmed by the measurement of the GALT enzyme activity in erythrocytes and/or *GALT* gene analysis. Current policy is to treat patients if the GALT enzyme activity in erythrocytes is below 15% of the reference range and/or two pathogenic variations in the *GALT* gene are detected[16].

Dietary treatment is currently the only available treatment and consists of a lifelong lactose free, galactose-restricted diet. Milk and dairy products are always restricted considering the significant amount of galactose in these products[16]. Until recently the extent of galactose restriction from products such as fruit, vegetables and legumes varied worldwide and even within countries. The amount of galactose in these products has been assessed and is now considered negligible[16,18]. Multiple studies have confirmed that small amounts of exogenous galactose are safe and do not have an adverse effect on long-term outcome[19-21]. Thus, a diet that includes products that contain small amounts or traces of galactose has been implemented in the International guideline for the management of classical galactosemia which was published in 2017.

Disease course

With the ability to reliably diagnose galactosemia and the initiation of dietary galactose restriction, the management of CG was considered complete for many years. This changed after the publication of a growing number of reports on long-term complications observed in CG patients[22-26]. The finding that long-term complications such as a delayed motor- and speech development, cognitive impairment, movement disorders and primary ovarian insufficiency (POI) were even frequently observed in patients who were detected and treated early as a result of newborn- and family screening, was disappointing[17,27-29]. Since then, research mainly focusses on two unsolved issues in classical galactosemia:

1. The pathophysiology of long-term complications
2. The unexplained high variability in clinical outcome

The pathophysiology

Several mechanisms have been suggested to be involved in the pathogenesis of long-term complications in CG.

Firstly, the accumulation of toxic metabolites is presumed to cause damage to tissues and organs. As a result of the GALT deficiency, galactose, galactitol and Gal-1-P levels are highly elevated in blood and tissues of CG patients. Even though Gal-1-P in erythrocytes and urinary galactitol rapidly decline after the initiation of a galactose -restricted diet, both galactitol and Gal-1-P levels remain elevated in dietary compliant patients[30-33]. The persistent elevation of these toxic metabolites despite dietary galactose restriction results from the ongoing endogenous production of galactose. The production of endogenous galactose in galactosemia patients is considerable (up to 1 gram/day)

and exceeds the galactose intake of patients on galactose-restricted diets by far[34]. The endogenous galactose production has been reported to be age dependent with a several fold higher production rate in children when compared with adults[35,36]. Importantly, the intake of exogenous galactose does not seem to influence the endogenous production of galactose[37]. As Gal-1-P and galactitol are not correlated[33] and a majority of studies found no correlation between biochemical parameters and clinical outcome[17,25,32,38,39], the extent to which these metabolites are elevated does not seem to explain the variability in clinical outcome. There is an ongoing debate on whether Gal-1-P is the most toxic metabolite or whether both galactitol and Gal-1-P contribute to the long-term complications observed in CG patients. The theory that Gal-1-P is the major factor arises from the finding of cataract as the only long-term complication in GALK deficient patients. These patients accumulate galactose and galactitol, but not Gal-1-P. On the other hand, a study performed in GALT deficient mice demonstrated accumulation of Gal-1-P with a distribution similar to humans without evidence of galactose toxicity[40]. Since the galactitol levels in these mice were significantly lower compared with humans, it has been suggested that the combination of galactitol and Gal-1-P contribute to the long-term complications of CG patients. As this theory is based on a rodent model and galactose toxicity and pathological processes may differ in humans[41], the development of animal models that mimic human phenotypes both biochemically and clinically may improve our knowledge on the pathophysiology of CG.

Secondly, in addition to the accumulation of metabolites prior to their conversion by GALT, a reduced production of metabolites downstream of the deficient GALT enzyme may contribute to the pathology of CG. Abnormalities in UDP sugars have been reported[42-45], but this was not confirmed in all tissues[46]. Since UDP-sugars are essential in the biosynthesis of glycoproteins and glycolipids, a reduced level of these UDP-sugars may lead to the abnormal glycosylation of proteins and lipids[41,42]. Indeed, glycosylation abnormalities have been demonstrated in CG patients. Remarkably, abnormalities were demonstrated both after galactose intoxication and over restriction of galactose[47-49]. Even though the extent of abnormalities in UDP sugars and glycosylation varies between patients and overlaps with controls[47-51], a correlation with long-term outcome has been suggested[52].

Thirdly, Gal-1-P itself has been demonstrated to directly affect glycosylation. Pathological concentrations of Gal-1-P have been shown to reduce UDP-glucose synthesis leading to decreased concentrations of both UDP-glucose and UDP-galactose, which is presumed to account for the glycosylation defects observed in patients[45].

In conclusion, both the presence of toxic metabolites (Gal-1-P and galactitol) and a shortage of products downstream of the deficient GALT enzyme resulting in abnormal glycosylation of proteins and lipids, are considered to contribute to the complications observed in CG patients.

The timing of (tissue) damage

The question remains whether damage occurs in utero or if pathological processes continue throughout life. The answer to this question is highly relevant in the development of new treatment options. If CG is indeed a degenerative disease, (early) treatment will improve the clinical outcome of patients whereas early/prenatal damage may be irreversible. An early initiation of a galactose-restricted diet as a result of NBS and family screening does not seem to prevent long-term complications. The finding that (some) younger siblings who started the galactose-restricted diet directly after birth may even demonstrate a worse clinical outcome than the later treated older sibling was unexpected[17,42]. As the galactose metabolism pathway develops around the tenth week of gestation and abnormal levels of metabolites have been found in fetuses from 20 weeks of gestation[53] some complications may already occur in utero. The finding that the endogenous production of galactose continues in spite of galactose restriction from the diet suggests ongoing damage[34]. If long-term complications result from ongoing pathological processes only, clinical outcome may be modifiable if these processes are elucidated and modification of these processes are possible.

The clinical outcome

In contrast to the neonatal crisis with symptoms affecting multiple organ systems, the long-term complications observed in patients mainly involve the brain with abnormalities in motor- language and speech development, cognitive impairment and movement disorders. Besides the brain, the ovaries are affected resulting in primary ovarian insufficiency (POI) in the majority of females[17,27-29].

The clinical outcome of patients is highly variable and cannot be explained by a clear genotype-phenotype correlation[54]. There are a few exceptions in which certain *GALT* gene variations provide residual GALT enzyme activity resulting in a milder and even normal clinical outcome.

Since the introduction of CG in the NBS program, individuals with Duarte variant Galactosemia (DG) have been identified. In DG, one GALT allele is severely impaired and the second *GALT* allele is partially impaired, resulting in residual GALT enzyme activity of 14–25%[42]. As their galactose oxidation capacity is comparable to healthy controls[55,56] and newborns with DG do not present with clinical symptoms, these individuals are not considered as patients and are not treated with a galactose-restricted diet in Europe[16]. In the USA, however, some metabolic centers do treat these patients as literature reporting on the long-term outcome of DG is scarce.

Another special group of patients are the patients with the homozygous p.Ser135Leu variation, with residual GALT enzyme activity in tissues other than erythrocytes. These patients have a different biochemical profile and may demonstrate a milder clinical outcome[57].

Since CG was included in the Dutch NBS in 2007, patients with previously unreported

genotypes and phenotypes have been identified[13]. These patients demonstrate residual GALT enzyme activity in erythrocytes up to 10% and have a different biochemical profile with undetectable Gal-1-P levels on a galactose-restricted diet. Currently, these patients are treated with a galactose-restricted diet as they are still young and their long-term outcome has to be awaited. The question remains if their residual enzyme activity is sufficient to provide a normal clinical outcome with dietary treatment and perhaps even without dietary treatment.

Despite these exceptions, most CG patients have classical disease-causing variations in the *GALT* gene resulting in a severe GALT enzyme deficiency with GALT enzyme activity in erythrocytes in the very low range (<3.3%) and a barely detectable galactose oxidation capacity (<2%)[55,56]. Interestingly, these patients can be found on opposite ends of the clinical outcome spectrum and even siblings with identical *GALT* gene variations demonstrate a high variability in clinical outcome, which is poorly understood[17,42].

To summarize, classical galactosemia is a rare inborn error of galactose metabolism. Despite early diagnosis and treatment, patients are still at risk to develop long-term complications. The pathogenesis of complications and the variability in clinical outcome of CG patients is poorly understood. Markers that are able to predict clinical outcome are urgently needed to provide an early prognosis, not only for classical phenotype patients in which the clinical outcome may vary substantially, but also for the variant patients detected by NBS with possible better outcomes. The uncertainty of the long-term outcome of a newborn CG patient is a severe burden on parents. Considering the large variability in clinical outcome, an individual evaluation and follow-up is currently advised in all patients[16]. An early prognosis will provide clarity for parents and patients and will enable the identification of patients in need for support. Moreover, prognostic markers may lead to individual dietary treatment to prevent both galactose intoxication and over restriction.

Towards individual prognostication

The question remains whether minor differences in residual GALT enzyme activity in different tissues could explain the differences in clinical outcome observed in patients. Even a slightly higher residual enzyme activity may lead to lower Gal-1-P levels resulting in less abnormal galactosylation and possibly a more favorable clinical outcome. Currently, the measurement of residual GALT enzyme activity in erythrocytes is the golden standard for diagnosis, but the method used is not able to reliably detect differences in enzyme activity below 3.3% (<1.1 μmol/h.g Hb)[58]. Therefore, new methods that are able to study the correlation between residual enzyme activity and clinical outcome are needed.

In this thesis, we studied the presence of long-term complications in patients with varying genotypes and phenotypes and investigated the association between clinical outcome and potential biochemical markers and residual galactose oxidation capacity.

In an attempt to gain knowledge on the pathophysiology of long-term complications in CG that originate from the brain, we used quantitative MRI techniques to investigate structural changes of the brain and normal appearing gray and white matter and studied its association with clinical outcome.

THESIS OUTLINE

In **Part 1** of this thesis, the spectrum of clinical outcome of CG patients is studied in more depth. In **Chapter 2**, we investigated the long-term complications in our cohort of patients and the association with two possible biochemical markers (Gal-1-P and *N*-glycans) and brain abnormalities on MRI.
In **Chapter 3**, we performed a systematic review on the cognitive functioning of patients with CG. In **Chapter 4**, we studied the neuropsychological and psychosocial functioning of patients in our cohort of patients with the use of a comprehensive neuropsychological assessment and the use of self- and proxy reported questionnaires.
Currently, we are not able to predict clinical outcome in patients, which is a major burden on patients and parents and hampers individualized prognostication and treatment. In **Part 2**, the results of the in vitro (**Chapter 5**) and in vivo (**Chapter 6**) galactose oxidation tests are presented. Our aim was to investigate if the presence of residual galactose oxidation capacity provides a more favorable clinical outcome by comparing the results of both tests between patients with a poor and normal clinical outcome. In **Chapter 5**, a previously developed method, galactose metabolite profiling (GMP), was used to determine residual galactose oxidation capacity in cultured skin fibroblasts of patients with the use of radioisotope labeled galactose. In **Chapter 6**, radioisotope labeled galactose was given to patients and (whole body) galactose oxidation capacity was determined by measuring the labeled galactose in exhaled air over time.
In **Chapter 7**, a more quantitative analysis of the brain MRIs of patients was performed to investigate both white and gray matter and if structural changes on MRI were associated with clinical outcome.
To explain complicated inborn errors of metabolism, including CG, to children and adolescents and to improve knowledge on disease and treatment, newly developed patient education materials were tested in **Chapter 8** (**Part 3**).
To conclude, a general discussion is provided in **Chapter 9**, which also includes the future perspectives for research on CG. An English and Dutch summary of this thesis are provided in **Chapter 10**.

REFERENCES

1. Costa S, Natália, Rossi P, Maldonado R. Evaluation of Lactose in Milk and Dairy Products. *International journal for innovation education and research*. 2013;1.
2. De Bruyn F, Beauprez J, Maertens J, Soetaert W, De Mey M. Unraveling the Leloir pathway of Bifidobacterium bifidum: significance of the uridylyltransferases. *Applied and environmental microbiology*. 2013;79(22):7028-35.
3. Holden HM, Rayment I, Thoden JB. Structure and function of enzymes of the Leloir pathway for galactose metabolism. *Journal of biological chemistry*. 2003;278(45):43885-8.
4. https://www.omim.org/entry/230400#63.
5. Calderon FR, Phansalkar AR, Crockett DK, Miller M, Mao R. Mutation database for the galactose-1-phosphate uridyltransferase (GALT) gene. *Human mutation*. 2007;28(10):939-43.
6. https://arup.utah.edu/database/GALT/GALT_display.php.
7. Von Reuss A. Zuckerausscheidung im Sauglingsalter. *Wiener Medizinische Wochenschrift*. 1908(58):799–801.
8. Isselbacher KJ, Anderson EP, Kurahashi K, Kalckar HM. Congenital galactosemia, a single enzymatic block in galactose metabolism. *Science*. 1956;123(3198):635-6.
9. Leloir L. Two Decades of Research on the Biosynthesis of Saccharides. Nobel Lecture. Amsterdam: Elsevier Publishing Company; 1970.
10. Pyhtila BM, Shaw KA, Neumann SE, Fridovich-Keil JL. Newborn screening for galactosemia in the United States: looking back, looking around, and looking ahead. *JIMD reports*. 2015;15:79-93.
11. Levy HL, Hammersen G. Newborn screening for galactosemia and other galactose metabolic defects. *Journal of pediatrics*. 1978;92(6):871-7.
12. Jumbo-Lucioni PP, Garber K, Kiel J, Baric I, Berry GT, Bosch A, et al. Diversity of approaches to classic galactosemia around the world: a comparison of diagnosis, intervention, and outcomes. *Journal of inherited metabolic disease*. 2012;35(6):1037-49.
13. Welling L, Boelen A, Derks TG, Schielen PC, de Vries M, Williams M, et al. Nine years of newborn screening for classical galactosemia in the Netherlands: Effectiveness of screening methods, and identification of patients with previously unreported phenotypes. *Molecular genetics and metabolism*. 2017;120(3):223-8.
14. https://ddrmd.nl/general/downloads/.
15. https://www.orpha.net/consor/cgi-bin/Clinics.php?lng=NL.
16. Welling L, Bernstein LE, Berry GT, Burlina AB, Eyskens F, Gautschi M, et al. International clinical guideline for the management of classical galactosemia: diagnosis, treatment, and follow-up. *Journal of inherited metabolic disease*. 2017;40(2):171-6.
17. Hughes J, Ryan S, Lambert D, Geoghegan O, Clark A, Rogers Y, et al. Outcomes of siblings with classical galactosemia. *Journal of pediatrics*. 2009;154(5):721-6.
18. Van Calcar SC, Bernstein LE, Rohr FJ, Yannicelli S, Berry GT, Scaman CH. Galactose content of legumes, caseinates, and some hard cheeses: implications for diet treatment of classic galactosemia. *Journal of agricultural and food chemistry*. 2014;62(6):1397-402.

19. Bosch AM, Bakker HD, Wenniger-Prick LJ, Wanders RJ, Wijburg FA. High tolerance for oral galactose in classical galactosaemia: dietary implications. *Archives of disease in childhood*. 2004;89(11):1034-6.
20. Van Calcar SC, Bernstein LE, Rohr FJ, Scaman CH, Yannicelli S, Berry GT. A re-evaluation of life-long severe galactose restriction for the nutrition management of classic galactosemia. *Molecular genetics and metabolism*. 2014;112(3):191-7.
21. Krabbi K, Uudelepp ML, Joost K, Zordania R, Ounap K. Long-term complications in Estonian galactosemia patients with a less strict lactose-free diet and metabolic control. *Molecular genetics and metabolism*. 2011;103(3):249-53.
22. Komrower GM, Lee DH. Long-term follow-up of galactosaemia. *Archives of disease in childhood*. 1970;45(241):367-73.
23. Kaufman FR, Kogut MD, Donnell GN, Goebelsmann U, March C, Koch R. Hypergonadotropic hypogonadism in female patients with galactosemia. *New England journal of medicine*. 1981;304(17):994-8.
24. Waisbren SE, Norman TR, Schnell RR, Levy HL. Speech and language deficits in early-treated children with galactosemia. *Journal of pediatrics*. 1983;102(1):75-7.
25. Waggoner DD, Buist NR, Donnell GN. Long-term prognosis in galactosaemia: results of a survey of 350 cases. *Journal of inherited metabolic disease*. 1990;13(6):802-18.
26. Schweitzer S, Shin Y, Jakobs C, Brodehl J. Long-term outcome in 134 patients with galactosaemia. *European journal of pediatrics*. 1993;152(1):36-43.
27. Bosch AM. Classical galactosaemia revisited. *Journal of inherited metabolic disease*. 2006;29(4):516-25.
28. Waisbren SE, Potter NL, Gordon CM, Green RC, Greenstein P, Gubbels CS, et al. The adult galactosemic phenotype. *Journal of inherited metabolic disease*. 2012;35(2):279-86.
29. Coss KP, Doran PP, Owoeye C, Codd MB, Hamid N, Mayne PD, et al. Classical Galactosaemia in Ireland: incidence, complications and outcomes of treatment. *Journal of inherited metabolic disease*. 2013;36(1):21-7.
30. Gitzelmann R. Galactose-1-phosphate in the pathophysiology of galactosemia. *European journal of pediatrics*. 1995;154(7 Suppl 2):S45-9.
31. Jakobs C, Schweitzer S, Dorland B. Galactitol in galactosemia. *European journal of pediatrics*. 1995;154(7 Suppl 2):S50-2.
32. Walter JH, Collins JE, Leonard JV. Recommendations for the management of galactosaemia. UK Galactosaemia Steering Group. *Archives of disease in childhood*. 1999;80(1):93-6.
33. Hutchesson AC, Murdoch-Davis C, Green A, Preece MA, Allen J, Holton JB, et al. Biochemical monitoring of treatment for galactosaemia: biological variability in metabolite concentrations. *Journal of inherited metabolic disease*. 1999;22(2):139-48.
34. Berry GT, Nissim I, Lin Z, Mazur AT, Gibson JB, Segal S. Endogenous synthesis of galactose in normal men and patients with hereditary galactosaemia. *The Lancet*. 1995;346(8982):1073-4.
35. Schadewaldt P, Kamalanathan L, Hammen HW, Wendel U. Age dependence of endogenous galactose formation in Q188R homozygous galactosemic patients. *Molecular genetics and metabolism*. 2004;81(1):31-44.

36. Berry GT, Moate PJ, Reynolds RA, Yager CT, Ning C, Boston RC, et al. The rate of de novo galactose synthesis in patients with galactose-1-phosphate uridyltransferase deficiency. *Molecular genetics and metabolism*. 2004;81(1):22-30.
37. Huidekoper HH, Bosch AM, van der Crabben SN, Sauerwein HP, Ackermans MT, Wijburg FA. Short-term exogenous galactose supplementation does not influence rate of appearance of galactose in patients with classical galactosemia. *Molecular genetics and metabolism*. 2005;84(3):265-72.
38. Schweitzer-Krantz S. Early diagnosis of inherited metabolic disorders towards improving outcome: the controversial issue of galactosaemia. *European journal of pediatrics*. 2003;162 Suppl 1:S50-3.
39. Shield JP, Wadsworth EJ, MacDonald A, Stephenson A, Tyfield L, Holton JB, et al. The relationship of genotype to cognitive outcome in galactosaemia. *Archives of disease in childhood*. 2000;83(3):248-50.
40. Ning C, Reynolds R, Chen J, Yager C, Berry GT, Leslie N, et al. Galactose metabolism in mice with galactose-1-phosphate uridyltransferase deficiency: sucklings and 7-week-old animals fed a high-galactose diet. *Molecular genetics and metabolism*. 2001;72(4):306-15.
41. Lai K, Tang M, Yin X, Klapper H, Wierenga K, Elsas L. ARHI: A new target of galactose toxicity in Classic Galactosemia. *Bioscience hypotheses*. 2008;1(5):263-71.
42. Fridovich-Keil JL, Walter JH. Part 7: Carbohydrates, Chapter 72: Galactosemia. The Online Metabolic and Molecular Bases of Inherited Disease, OMMBID; Valle D.L., Antonarakis S, Ballabio A, Beaudet A.L., Mitchell G.A.(Eds.). McGraw Hill, New York.
43. Ng WG, Xu YK, Kaufman FR, Donnell GN. Deficit of uridine diphosphate galactose in galactosaemia. *Journal of inherited metabolic disease*. 1989;12(3):257-66.
44. Keevill NJ, Holton JB, Allen JT. The investigation of UDPGlucose and UDPGalactose concentration in red blood cells of patients with classical galactosaemia. *Clinica chimica acta*. 1993;221(1-2):135-42.
45. Lai K, Langley SD, Khwaja FW, Schmitt EW, Elsas LJ. GALT deficiency causes UDP-hexose deficit in human galactosemic cells. *Glycobiology*. 2003;13(4):285-94.
46. Keevill NJ, Holton JB, Allen JT. UDP-glucose and UDP-galactose concentrations in cultured skin fibroblasts of patients with classical galactosaemia. *Journal of inherited metabolic disease*. 1994;17(1):23-6.
47. Coss KP, Byrne JC, Coman DJ, Adamczyk B, Abrahams JL, Saldova R, et al. IgG *N*-glycans as potential biomarkers for determining galactose tolerance in Classical Galactosaemia. *Molecular genetics and metabolism*. 2012;105(2):212-20.
48. Coman DJ, Murray DW, Byrne JC, Rudd PM, Bagaglia PM, Doran PD, et al. Galactosemia, a single gene disorder with epigenetic consequences. *Pediatric research*. 2010;67(3):286-92.
49. Coss KP, Hawkes CP, Adamczyk B, Stockmann H, Crushell E, Saldova R, et al. *N*-glycan abnormalities in children with galactosemia. *Journal of proteome research*. 2014;13(2):385-94.
50. Liu Y, Xia B, Gleason TJ, Castaneda U, He M, Berry GT, et al. *N*- and O-linked glycosylation of total plasma glycoproteins in galactosemia. *Molecular genetics and metabolism*. 2012;106(4):442-54.
51. Gibson JB, Reynolds RA, Palmieri MJ, Berry GT, Elsas LJ, 2nd, Levy HL, et al. Comparison of erythrocyte uridine sugar nucleotide levels in normals, classic galactosemics, and patients with other metabolic disorders. *Metabolism*. 1995;44(5):597-604.

52. Maratha A, Stockmann H, Coss KP, Estela Rubio-Gozalbo M, Knerr I, Fitzgibbon M, et al. Classical galactosaemia: novel insights in IgG *N*-glycosylation and *N*-glycan biosynthesis. *European Journal of Human Genetics*. 2016;24(7):976-84.

53. Holton JB. Effects of galactosemia in utero. *European journal of pediatrics*. 1995;154(7 Suppl 2):S77-81.

54. Tyfield LA. Galactosaemia and allelic variation at the galactose-1-phosphate uridyltransferase gene: a complex relationship between genotype and phenotype. *European journal of pediatrics*. 2000;159 Suppl 3:S204-7.

55. Berry GT, Singh RH, Mazur AT, Guerrero N, Kennedy MJ, Chen J, et al. Galactose breath testing distinguishes variant and severe galactose-1-phosphate uridyltransferase genotypes. *Pediatric research*. 2000;48(3):323-8.

56. Berry GT, Nissim I, Mazur AT, Elsas LJ, Singh RH, Klein PD, et al. In vivo oxidation of [13C] galactose in patients with galactose-1-phosphate uridyltransferase deficiency. *Biochemical and molecular medicine*. 1995;56(2):158-65.

57. Lai K, Langley SD, Singh RH, Dembure PP, Hjelm LN, Elsas LJ, 2nd. A prevalent mutation for galactosemia among black Americans. *Journal of pediatrics*. 1996;128(1):89-95.

58. Shin-Buhring Y, Osang M, Ziegler R, Schaub J. A method for galactose-1-phosphate uridyltransferase assay and the separation of its isozymes by DEAE-cellulose column chromatography. *Clinica chimica acta*. 1976;70(3):371-7.

PART 1

THE SPECTRUM OF CLINICAL OUTCOME

Chapter 2

Deep phenotyping classical galactosemia: clinical outcomes and biochemical markers

M.M. Welsink-Karssies, S. Ferdinandusse, G.J. Geurtsen, C.E.M. Hollak, H.H. Huidekoper, M.C.H. Janssen, J.G. Langendonk, J.H. van der Lee, R. O'Flaherty, K.J. Oostrom, S.D. Roosendaal, M.E. Rubio-Gozalbo, R. Saldova, E.P. Treacy, F.M. Vaz, M.C. de Vries, M. Engelen*, A.M. Bosch*

* Authors contributed equally to this manuscript

Brain Communications. 2020 Jan;2 (1).

ABSTRACT

Early diagnosis and dietary treatment do not prevent long-term complications, which mostly affect the central nervous system in classical galactosemia patients. The clinical outcome of patients is highly variable and there is an urgent need for prognostic biomarkers. The aim of this study was first to increase knowledge on the natural history of classical galactosemia by studying a cohort of patients with varying geno- and phenotypes and second to study the association between clinical outcomes and two possible prognostic biomarkers. In addition, the association between abnormalities on brain MRI and clinical outcomes was investigated.

Classical galactosemia patients visiting the galactosemia expertise outpatient clinic of the Amsterdam University Medical Center were evaluated according to the International classical galactosemia guideline with the addition of an examination by a neurologist, serum immunoglobulin G *N*-glycan profiling and a brain MRI. The biomarkers of interest were galactose-1-phosphate levels and *N*-glycan profiles, and the clinical outcomes studied were intellectual outcome (IQ), and the presence or absence of movement disorders and/or primary ovarian insufficiency.

Data of 56 classical galactosemia patients are reported. The IQ ranged from 45 to 103 (mean 78 ± 14) and was below 85 in 62%. Movement disorders were found in 17 (47%) out of 36 tested patients. In females aged 12 years and older, primary ovarian insufficiency was diagnosed in 12 (71%) out of 17 patients. Significant differences in *N*-glycan peaks were found between controls and patients. However, no significant differences in either *N*-glycans or galactose-1-phosphate levels were found between patients with a poor intellectual outcome (IQ<85) and normal intellectual outcome (IQ≥85), and with or without movement disorders or primary ovarian insufficiency. The variant patients detected by newborn screening, with previously unknown geno- and phenotypes and currently no long-term complications, demonstrated significantly lower galactose-1-phospate levels than classical patients (p<0.0005). Qualitative analysis of the MRIs demonstrated brain abnormalities in 18 out of 21 patients, more severely in patients with a lower intellectual outcome and/or with movement disorders.

This study demonstrates a large variability in clinical outcome, which varies from a below average intelligence, movement disorders and in females primary ovarian insufficiency to a normal clinical outcome. In our cohort of classical galactosemia patients, galactose-1-phosphate levels and *N*-glycan variations were not associated with clinical outcomes, but galactose-1-phosphate levels did differentiate between classical and variant patients detected by newborn screening. The correlation between brain abnormalities and clinical outcome should be further investigated by quantitative analysis of the MR images. The variability in clinical outcome necessitates individual and standardized evaluation of all classical galactosemia patients.

INTRODUCTION

Classical Galactosemia (CG, OMIM 230400) is one of the more frequent inborn errors of metabolism caused by a severe deficiency of the enzyme galactose-1-phosphate uridylyltransferase (GALT, EC 2.7.7.12). In newborns with CG, the ingestion of galactose causes life-threatening illness. In the Netherlands, CG was implemented in the newborn screening (NBS) program in 2007. Hereafter, 31 patients have been identified: 25 patients with a classical phenotype and 6 patients with previously unreported genotypes and phenotypes. Based on these findings, the incidence of CG is estimated to be 1:52.800 in the Netherlands[1]. Despite an early diagnosis by NBS and treatment with a galactose-restricted diet, patients are at risk to develop long-term complications of the central nervous system, such as abnormalities in motor and speech development, cognitive impairment, movement disorders and (MDs), and ovarian insufficiency in females[2-6]. There is a broad spectrum of clinical manifestations ranging from fully normal to severely impaired, even within families with identical mutations[4,7]. Unfortunately, at this time, the clinical outcome of individual patients cannot be predicted because prognostic biomarkers are lacking. This is a severe burden on parents and patients and hampers the development of new therapeutic options.

The pathophysiology of long-term complications is poorly understood. The endogenous production of significant amounts of galactose causes a persistent elevation of harmful metabolites, such as galactose-1-phosphate (Gal-1-P), which is considered a major factor [7-9]. Gal-1-P has also been demonstrated to affect galactosylation[10], and *N*-glycan abnormalities are reported in CG patients [7,11,12]. Remarkably, galactosylation abnormalities were seen both after galactose intoxication and galactose over restriction and improvement in galactosylation patterns after moderate dietary galactose relaxation could indicate that dysregulated glycosylation pathways are modifiable[11-13]. We hypothesize that individual differences in the extent of galactose intoxication and galactosylation abnormalities may contribute to the variability of the clinical outcome spectrum observed in CG patients. Previous studies suggested two possible predictors of clinical outcome. First, an association between Gal-1-P levels and long-term outcome in general and verbal dyspraxia in particular has been found [14-16]. Second, differences in glycosylation patterns of immunoglobulin G (IgG) *N*-glycans were found between patients and controls [12,17-19] linked to specific glycan synthesis gene abnormalities with a proposed correlation with intellectual outcome[19]. Also, abnormal IgG *N*-glycosylation, as well as both inflammatory and glycosylation gene expression in CG patients, has been correlated with fertility endocrine markers in female CG patients with ovarian insufficiency[20].

The aim of this study was first to gain knowledge on the natural history of CG by performing deep phenotyping in a relatively large cohort of patients, who are assessed according to the International guideline[21] and second to study the association between two possible predictors of clinical outcome, Gal-1-P levels and IgG *N*-glycan abnormalities,

and long-term complications. The effect of genotype and an early initiation of treatment were also investigated, as well as the association between abnormalities on MRI and clinical outcome.

METHODS

Study design and recruitment

In this cohort study, clinical data were retrieved from the medical records of CG patients who visited the multidisciplinary galactosemia outpatient clinic in the Amsterdam University Medical Center (Amsterdam UMC) or who are treated in other metabolic centers and participated in research in the Amsterdam UMC. All included patients consented to the use of their data for research purposes. Data on intellectual and neurological functioning were collected prospectively based on a predetermined standardized assessment. Serum samples, stored in the Amsterdam UMC Galactosemia Biobank after informed consent of patients and/or parents, were used to measure IgG *N*-glycans. The MRI was performed for research purposes, and informed consent was obtained from all included patients. The use of the serum samples and the MRI study was approved by the local institutional review board, and a waiver was given for the data collection.

Data storage

All data were stored in an electronic clinical report form in Castor EDC, a good clinical practice compliant data management system[22].

Inclusion and exclusion criteria

Patients with an erythrocyte GALT activity <15% of the reference mean[23] and/or two known pathogenic variations in the *GALT* gene were eligible for participation in this study[21]. Patients with a second genetic diagnosis influencing clinical outcome were excluded.

Patient groups

For analyses, patients were divided into three groups:

- Classical phenotypes: patients with two pathogenic *GALT* gene mutations with an erythrocyte GALT enzyme activity below the limit of quantitation of the enzyme assay (<3.3%;<1.1 μmol/h.g Hb). Patients with a classical phenotype, hereafter classical patients, can be divided into screened patients (detected after NBS or family screening (FS) because of an older sibling with CG with dietary treatment started immediately after birth) and non-screened patients (diagnosed before the implementation of NBS, after a clinical presentation with CG-related symptoms).
- Variant patients: patients detected since the introduction of NBS in 2007[1]. These patients with previously unknown genotypes and phenotypes were asymptomatic

at the time of diagnosis and have residual GALT enzyme activities up to 10% in erythrocytes and erythrocyte Gal-1-P levels below the detection limit (<0.05 μmol/g Hb) under treatment.

- Homozygous p.Ser135Leu patients: patients with GALT deficiency in erythrocytes but residual GALT enzyme activity in other tissues that may improve clinical outcome[24].

Clinical outcome measures

Patients were assessed according to the International CG guideline[21]. In patients who received a standardized age-specific intelligence test, the IQ was used as derivative of intelligence[25]. The used intelligence tests were the Bayley Scales of Infant and Toddler Development, the Wechsler Preschool and Primary Scale of Intelligence, the Wechsler Intelligence Scale for Children and the Wechsler Adult Intelligence Scale. A poor intellectual outcome was defined as an IQ<85 and a normal intellectual outcome was defined as an IQ≥85. Adult patients of whom no IQ was available and who were unable to live independently due to cognitive impairment were considered to have a poor intellectual outcome.

The neurological examination to assess the presence or absence of MDs and standardized screening for tremors (Fahn–Tolosa–Marin Clinical Rating Scale for Tremor, scores ranging from 0 to 84)[26] and dystonia (Fahn–Marsden Rating Scale, scores ranging from 0 to 120)[27] were performed by one (pediatric) neurologist (M.E.) and documented for all patients. Since the rating scales are not validated in children, scores of pediatric patients below 12 years were not reported in this article.

Data on speech development and the development of gross and fine motor skills reported by the treating physician were retrieved from the medical records.

Information on the presence or absence of primary ovarian insufficiency (POI) was retrieved from the medical records of female patients.

Gal-1-P measurement in erythrocytes

Before studying the association between Gal-1-P and clinical outcomes, we evaluated the stability of erythrocyte Gal-1-P (hereafter Gal-1-P) in our cohort of classical patients. We found relatively stable Gal-1-P levels from the age of 12 months (**Supplementary Figure 1**) and therefore, measurements performed before the age of 12 months were excluded. The most recent Gal-1-P levels reported in this article were measured by gas chromatography-mass spectrometry and were <0.82 μmol/g Hb in diet adherent patients. To evaluate Gal-1-P as a prognostic biomarker, only dietary adherent patients were included. Considering the most recent Gal-1-P levels of the classical patients were within a narrow range, lifetime Gal-1-P was also evaluated in this subgroup.

The method to measure Gal-1-P changed from spectrophotometry[28] to gas chromatography-mass spectrometry in September 2016. Therefore, lifetime Gal-1-P was defined as the mean of all Gal-1-P levels measured between 12 months of age until September 2016.

Patients with a minimum of five measurements were included in the analysis. The Gal-1-P levels measured by spectrophotometry were <0.58 μmol/g Hb in diet adherent patients.

IgG *N*-glycan measurement

Analysis of serum IgG *N*-glycans was performed by the National Institute for Bioprocessing Research and Training (Dublin, Ireland) using automated glycan preparation method linked to ultra-performance liquid chromatography as previously described[29]. The percentage areas of 28 high-resolution IgG *N*-glycan peaks (GPs) were quantified in patients and healthy controls[18]. Based on previous research, main glycans were assigned and *N*-glycan features were calculated including agalactosylated (G0), monogalactosylated (G1) and digalactosylated (G2) structures, resulting G-ratios G0/G1, G0/G2 as well as core fucosylated bisected neutral glycans, afucosylated bisected neutral glycans and total bisected glycans[19,30,31]. The results of the adult CG patients were compared to the 79 adult control samples, previously reported by Maratha et al.[19], and the results of the pediatric CG patients in our cohort were compared to nine pediatric control samples, previously reported by Coss et al.[17].

MRI protocol

An open-bore 3.0-T MRI scanner (Tesla Philips Ingenia scanner) with a 32-channel head coil was used to perform MRI of the brain. The MRI protocol included three-dimensional T1-weighted and fluid-attenuated inversion recovery (FLAIR) sequences with isotropic voxels.

Image analysis

All brain MR images were scored by one neuro-radiologist (S.D.R.) blinded for the clinical outcome of patients. The MR images were evaluated for focal white matter abnormalities and abnormalities of the basal ganglia, thalamus and cortex. The Fazekas scale was used to quantify white matter abnormalities. Cerebellar atrophy was assessed in the cerebellum and vermis. Cerebral atrophy was scored with the use of the global cortical atrophy (GCA) scale. White matter abnormalities and atrophy were scored if the MR images deviated from the existing age-based standards.

Statistical analysis

SPSS version 25 (SPSS Inc. Chicago, IL, USA) was used to perform all statistical analyses. Median and ranges were presented since data followed a non-normal distribution. Descriptive statistics of patients and MR images were reported. To determine if there were statistically significant differences in continuous variables and proportions between two groups, the Mann-Whitney U test and Chi-square tests (or Fisher's exact test) were used, respectively. Considering the differences between the *N*-glycans profiles of children and adults[32], both groups were analyzed separately. The Spearman's rank coefficient test

was used to test for associations, and in case of a significant association, linear regression or logistic regression was used where appropriate to test for correlations. P-values <0.05 were considered statistically significant. If multiple tests were carried out regarding a single hypothesis, the results were corrected using the Bonferroni-Holm method.

RESULTS

In total, 70 out of 77 CG patients visiting the Amsterdam UMC consented to the use of their clinical data. Four patients were excluded because of a second diagnosis influencing clinical outcome and 10 patients were excluded because of missing data. Demographics and clinical outcome data of 56 patients are presented in **Table 1** on an individual level and in **Table 2** on a group level.

Previously unreported mutations detected since the implementation of NBS[1] were found in 7 out of 55 patients, and the homozygous p.Ser135Leu mutation was found in 2 out of 55 patients (**Table 1**). In 9 out of 54 patients, erythrocyte GALT enzyme activity was above the limit of quantitation of the enzyme assay (>3.3%; >1.1 μmol/h.g Hb) and ranged from 3.6% to 9.3%.

Diagnosis and dietary treatment

Data on diagnosis were available in 50 patients. Twenty-five patients were diagnosed in the first six weeks of life because of CG-related symptoms, 12 patients were diagnosed after NBS and 10 patients were diagnosed after family screening (FS). In 9 out of 10 FS diagnosed patients, no CG-related symptoms were reported in the newborn period and data were missing in one patient. Three patients were diagnosed late (**Table 1**): two homozygous p.Ser135Leu patients were diagnosed at age 7 months and 10 years because of feeding difficulties and cataract and because of visual impairment due to cataract, respectively. The third patient was diagnosed at age 9 months after a diagnostic delay and information regarding CG-related symptoms is missing.

In 32 out of 50 patients, of whom 7 patients were diagnosed by NBS, CG-related symptoms were reported at diagnosis. The most frequently reported symptoms were jaundice (81%), elevated liver enzymes (58%), clinical suspicion of sepsis (52%, positive blood culture 35%), feeding difficulties (45%), vomiting (33%) and coagulopathy (30%). In the non-screened patients diagnosed because of critical illness (with the exclusion of the three late diagnosed patients), the lactose free, galactose-restricted diet was started at a median age of 10 days (4–39) with the confirmation of diagnosis at a median age of 11 days (4–46). In patients detected by NBS, the diet was started at a median age of 7 days (5–8) with the confirmation of diagnosis at a median age of 8 days (6–11).

In 52 out of 56 patients, strict dietary adherence was reported, while 4 out of 56 patients did not adhere to the diet in the past or at the most recent visit.

Table 1. Overview of included patients, demographics and clinical outcome

Pt ID	Gender	Group	GALT_1 / GALT_2	GALT activity, %	IQ	Tremor/ TRS	Dystonia / DRS	POI	BMD Z-score ≤ -2 SD
1	M	non-screened	p.Gln188Arg / 400delT	<3.3	-	-	-	NA	No
2	F	non-screened	p.Gln188Arg / p.Lys285Asn	<3.3	77	-	-	No	-
3	M	non-screened	p.Gln188Arg / p.Ser135Trp	<3.3	>85	-	-	NA	No
4	M	non-screened	- / -	<3.3	-	-	-	NA	No
5	F	non-screened	p.Ser135Trp / p.Arg51Gln	<3.3	78	-	-	Yes	Yes
6	M	non-screened	p.Gln188Arg / p.Gln188Arg	<3.3	-	Yes / 7	No / 0	NA	-
7	F	non-screened	p.Gln188Arg / p.Gln188Arg	<3.3	78	-	-	Yes	No
8	F	non-screened	p.Gln188Arg / p.Gln188Arg	<3.3	-	Yes / -	No / -	Yes	No
9	M	non-screened	p.Gln188Arg / p.Lys127Glu	<3.3	70	-	-	NA	-
10	M	non-screened	p.Gln188Arg / p.Gln188Arg	<3.3	74	-	-	NA	No
11	M	non-screened	p.Gln188Arg / p.Gln188Arg	<3.3	45	Yes / 38	Yes / 6	NA	No
12d	F	non-screened	p.Gln188Arg / p.Gln188Arg	<3.3	81	Yes / 4	No / 0	Yes	No
13	M	non-screened	c.329-2A>C/ c.329-2A>C	<3.3	-	Yes / -	No / -	NA	Yes
14	F	non-screened	p.Gln188Arg / p.Gln188Arg	<3.3	83	-	-	Yes	No
15	F	non-screened	p.Gln188Arg / p.Ser135Trp	<3.3	57	-	-	Yes	No
16	F	non-screened	p.Gln188Arg / p.Gln188Arg	-	53	No / 4	No / 0	Yes	No
17	F	non-screened	c.329-2A>C/ c.329-2A>C	<3.3	-	No / -	No / -	Yes	No
18	F	non-screened	p.Ser135Leu / p.Ser135Leu	<3.3	71	No / 1	No / 0	No	-
19	F	non-screened	p.Gln188Arg / p.Gln188Arg	<3.3	82	No/ 0	No / 0	?	No
20c	F	non-screened	p.Gln188Arg / p.Leu195Pro	<3.3	88	Yes / 5	No / 0	Yes	No
21	F	non-screened	p.Gln188Arg / p.Gln188Arg	<3.3	71	-	-	?	No
22	M	non-screened	p.Gln188Arg / p.Gln188Arg	-	91	-	-	NA	No
23	M	non-screened	p.Gln188Arg / p.Lys195Pro	<3.3	93	No / 0	No / 0	NA	No
24	M	non-screened	p.Gln188Arg / p.Ser135Trp	<3.3	98	No / 0	No / 0	NA	No
25	F	non-screened	p.Arg205* / p.Trp316*	<3.3	97	-	-	No	Yes
26a	M	non-screened	p.Gln188Arg / p.Lys285Asn	<3.3	76	Yes / 7	No / 0	NA	No
27a	M	non-screened	p.Gln188Arg / p.Lys285Asn	<3.3	86	Yes / 6	No / 0	NA	No
28	F	non-screened	p.Ser135Leu / p.Ser135Leu	3,9	61	No / 0	No / 0	No	No
29	M	non-screened	p.Gln188Arg / p.Ser135Trp	3,9	65	Yes / 6	No / 0	NA	No

30e	M	non-screened	p.Gln188Arg / c.377+7A>C	<3.3	87	No / 3	No / 0	NA	No
31	M	non-screened	p.Gln188Arg / p.Lys285Asn	<3.3	49	Yes / 29	Yes / 4	NA	No
32	M	non-screened	p.Gln188Arg / p.Leu195Pro	<3.3	103	Yes / 8	No / 0	NA	No
33b	F	non-screened	p.Arg148Gln / p.Trp316*	<3.3	88	Yes / 13	No / 0	?	No
34	F	non-screened	p.Gln188Arg / p.Gln188Arg	<3.3	46	Yes / -	No / -	-	No
35f	F	screened, FS	p.Gln188Arg / p.Gln188Arg	<3.3	77	Yes / 4	No / 0	Yes	No
36	M	screened, FS	p.Gln188Arg / p.Lys127Glu	<3.3	61	-	-	NA	No
37	F	screened, FS	p.Gln188Arg / p.Arg148Gln	<3.3	-	-	-	No	-
38f	F	screened, FS	p.Gln188Arg / p.Gln188Arg	<3.3	71	-	-	Yes	No
39d	F	screened, FS	p.Gln188Arg / p.Gln188Arg	<3.3	61	Yes / 34	Yes / 1	Yes	Yes
40c	F	screened, FS	p.Gln188Arg / p.Leu195Pro	<3.3	52	Yes / 10	Yes / 2	?	-
41e	M	screened, FS	p.Gln188Arg / c.377+7A>C	<3.3	95	-	-	NA	Yes
42b	M	screened, FS	p.Arg148Gln / p.Trp316*	<3.3	68	Yes / -	No / -	NA	No
43	F	screened, NBS	p.Gln188Arg / p.Leu195Pro	<3.3	82	No / -	No / -	-	-
44	F	screened, NBS	p.Gln188Arg / p.Gln188Arg	<3.3	64	No / -	No / -	-	-
45	M	screened, NBS	p.Gln188Arg / p.Gln188Arg	<3.3	89	No / -	No / -	NA	-
46	M	screened, NBS	p.Gln188Arg / p.Gln188Arg	<3.3	64	No / -	No / -	NA	-
47	F	screened, NBS	p.Gln188Arg / p.Gln188Arg	<3.3	70	No / -	No / -	-	-
48	M	screened, NBS	p.Ser135Leu / p.*380Argext*50	<3.3	95	No / -	No / -	NA	-
49	F	screened, NBS	p.Gln188Arg / p.Gln188Arg	<3.3	83	-	-	-	-
50g	F	Variant	p.Val128Ile / p.Val128Ile	8.7	86	No / -	No / -	-	-
51g	F	Variant	p.Val128Ile / p.Val128Ile	8.7	89	No / -	No / -	-	-
52	M	Variant	p.Gln188Arg / p.Met219Lys	7.2	96	No / -	No / -	NA	-
53	F	Variant	p.Gln188Arg / c.1-96T>G	3.6	86	No / -	No / -	-	-
54g	M	Variant	p.Val128Ile / p.Val128Ile	9.3	91	No / -	No / -	NA	-
55	M	Variant	p.Arg201His / p.Arg201His	8.9	-	-	-	NA	-
56g	F	Variant	p.Val128Ile / p.Val128Ile	6.3	-	-	-	-	-

Notes. Pt ID: patient ID, F: Female, M: Male, FS: family screening, NBS: newborn screening, IQ: intelligence quotient, TRS: tremor rating scale, DRS: dystonia rating scale, POI: primary ovarian insufficiency, BMD: bone mineral density. NA: not applicable, - : data not available, ? : data unclear, a,b,c,d,e,f: sibs, patients are ordered by group and listed by decreasing age, ▩: age below 12 years. Bold: patients diagnosed late.

Table 2. Patient groups and clinical outcomes

	n	All patients (*n*=56)	*n*	Classical patients (*n*=47)	*n*	Variant patients (*n*=7)	*n*	Homozygous p.Ser135Leu (*n*=2)
Gender	56		47					
- Male	26		23		3			
- Female	30		24		4		2	
Age (years)	56	18.0 (0 – 48)	47	21.0 (4 – 48)	7	5.0 (0 – 7)	2	16 / 23
- Pediatric patients (<18 years)	27		19		7		1	
- Adult patients (≥18 years)	29		28				1	
GALT-activity, %	54		45		7		2	
- <3.3	45		44				1	
- ≥3.3	9		1		7		1	
IQ	47	78 (45 – 103)	40	77 (45 – 103)	5	89 (86 – 96)	2	61 / 71
- IQ<85	29		27				2	
- IQ≥85	18		13		5			
Neurology	36		29		5		2	
- MDs, No	19		12		5		2	
- MDs, Yes	17		17					
TRS	20	5.5 (0 – 38)	18	6.0 (0 – 38)			2	0 / 1
DRS	20	0.0 (0 – 6)	18	0.0 (0 – 6)			2	0 / 0
Endocrinology (females ≥12 years)	21		19				2	
Puberty induced, No	11		9				2	
Puberty induced, Yes	8		8					
- POI, No	5		3				2	
- POI, Yes	12		12					
Gal-1-P (µmol/g Hb) in erythrocytes*	50	0.35 (0.0 – 0.70)	43	0.37 (0.16 – 0.70)	7	0.0 (0 – 0.13)	2	0.14 / 0.13
Galactitol (mmol/mol creatinine) in urine*	34	120 (0 – 311)	25	140 (109 – 311)	7	10 (0.0 – 97)	2	69 / 58

Notes. Data reported in median and ranges. IQ: intelligence quotient, MDs: movement disorders, TRS: tremor rating scale, DRS: dystonia rating scale, POI: primary ovarian insufficiency, Gal-1-P: galactose-1-phosphate. * most recent value is presented.

Deep phenotyping: the clinical outcome spectrum

Motor and speech development

The development of gross and fine motor skills was abnormal in 9 out of 43 (21%) and 5 out of 41 (12%) patients with available data, respectively. Speech and language development was abnormal in 25 out of 43 patients (58%), with a language delay in 11 (44%), a speech defect in 4 (16%) and a combination of both in 7 patients (28%). In three patients (12%), the speech and language development problems were not further specified. The frequency of abnormal motor, speech and language development did not differ between classical non-screened and classical screened patients.

Ophthalmology

In 12 out of 49 patients, cataract was present at the most recent ophthalmic evaluation, 11 out of 49 had cataract in the past which had resolved and 26 out of 49 patients have never been diagnosed with cataract. The most frequently reported description was

stable, nuclear cataract without visual impairment.

General intelligence

In 47 patients, IQ testing was performed. As a group, the included patients demonstrated an overall below average but highly variable intelligence with the IQ ranging between 45 and 103 (mean 78, 14 SD). Three adult patients were unable to live independently due to cognitive impairment. A total of 31 patients had a poor intellectual outcome, and 19 patients had a normal intellectual outcome. There was no significant difference in IQ between classical non-screened and classical screened patients (**Table 3**).

Neurology

In 36 out of 56 patients, an examination by a (pediatric) neurologist including a standardized screening for tremors and dystonia was performed at the outpatient clinic visit. Neurologic complaints were self-reported by 10 out of 36 patients (four adults) or their parents (six children) and ranged from a poor handwriting to symptoms interfering with activities in daily life. In 17 out of 36 (9/22 children, 8/14 adults), the neurologic examination was abnormal of whom 12 patients had a subtle tremor, 1 patient had an evident tremor and 4 patients demonstrated both an evident tremor and dystonia.

Out of all the patients with a subtle tremor, 9 had an action tremor, 2 had a postural and action tremor and 1 patient had a postural and intention tremor. Of the five patients with an evident tremor, two patients had a combined action and postural tremor and two patients had an action tremor only and in one patient the nature of the tremor was not specified.

Twenty patients aged 12 years and older underwent standardized screening for tremors and dystonia by the previously described rating scales (**Table 1**).

Neurologic comorbidity was reported in patients with and without movement disorders (MDs). The most severely affected patient had a complex MD with a dystonic tremor and spastic paraparesis. This patient had a meningitis in the neonatal period and suffers from epileptic seizures (tonic–clonic and focal with impaired consciousness) treated with multiple anti-epileptic drugs. Neurological comorbidity was reported in two other patients with an MD [epilepsy (n=1) and neonatal meningitis (n=1)] and in two patients without an MD [traumatic brain injury leading to a skull fracture with increased intracerebral pressure (n=1) and neonatal meningitis (n=1)].

In patients with an MD, the motor development was abnormal in 42% and the speech development was abnormal in 75%, which is in contrast to patients without an MD in whom motor and speech development were abnormal in 16% and 38%, respectively.

The frequency of MDs did not differ between classical non-screened and classical screened patients (**Table 3**).

Magnetic Resonance Imaging

A total of 21 CG patients (8–47 years, 9 males and 12 females) underwent MRI of the brain, and in 18 out of 21 patients, brain abnormalities were found (**Table 4**).

The supratentorial white matter appeared normal in 12 patients (median age 17 years, 8–31). In 9 out of 21 patients (median age 26 years, 18–47), focal deep white matter abnormalities were found, most frequently in the frontoparietal region, with confluent white matter abnormalities in one patient. In none of the patients, focal abnormalities were found in the infratentorial white matter, basal ganglia, thalamus and cortex.

In 4 patients no atrophy was detected, 12 patients demonstrated both cerebral and cerebellar atrophy, 4 patients demonstrated cerebral atrophy only and 1 patient demonstrated cerebellar atrophy only. Cerebral atrophy varied from minimal atrophy in one region of the brain to atrophy in multiple regions, up to a GCA score of 2 per region. Parietal (16/17) and frontal (9/17) were the most affected cerebral regions, and the vermis was the most affected cerebellar region (13/17).

All patients who underwent an MRI received IQ testing. The median IQ was 78 (49–98). Six out of 21 patients had a normal intellectual outcome (IQ≥85). Thirteen out of 21 patients were examined by a neurologist, who diagnosed an MD in eight patients. The presence of white matter abnormalities was not associated with IQ nor MDs. Patients with a poor intellectual outcome (IQ<85) had significantly more cerebral atrophy (higher GCA scores) than patients with a normal intellectual outcome (p=0.011). The IQ (as continuous measure) was associated with cerebral atrophy (the GCA score) (p=0.014). The presence of cerebellar atrophy was significantly higher in patients with a poor intellectual outcome (p=0.014), and the IQ (as a continuous measure) was associated with cerebellar atrophy (p=0.028). MDs were associated with GCA scores (p=0.041) and patients with MDs demonstrated significantly higher GCA scores than patients without MDs (p=0.048). Patients with MDs did not demonstrate a higher frequency of cerebellar atrophy.

The frequency of MRI abnormalities (white matter abnormalities, cerebral and cerebellar atrophy) did not differ between classical non-screened and classical screened patients.

Endocrinology

Of 21 out of 30 females aged 12 years and older at the time of data collection, data on the endocrinological outcome were available. During the most recent visit, the menarche had occurred in all patients. The median age at puberty induction (n=8) was 12 years (11–15 years) with the menarche at a median age of 15 years (12–17 years), while a spontaneous menarche (n=11) occurred at a median age of 13.5 years (12–16 years). The diagnosis POI was reported in 12 out of 17 females and is uncertain in four. All females with POI received hormone replacement therapy at the most recent visit. Both p.Ser135Leu homozygous patients had a normal puberty development and have no POI. The frequency of POI did not differ between classical non-screened and classical screened patients (**Table 3**).

Table 3. Clinical outcomes of non-screened and screened classical patients

	n	Classical non-screened patients (*n*=32)	*n*	Classical screened patients (*n*=15)	P-value
Gender	32		15		
- Male	17		6		0.534
- Female	15		9		
Age (years)	32	24.5 (11 – 48)	15	10.0 (4 – 32)	**0.002**
- Pediatric patients (<18 years)	9		10		**0.024**
- Adult patients (≥18 years)	23		5		
GALT-activity, %	30		15		
- <3.3	29		15		1.000
- ≥3.3	1				
IQ	26	78.0 (45 – 103)	14	71.0 (52 – 95)	0.421
- IQ<85	16		10		1.000
- IQ≥85	10		4		
Neurology	19		10		
- MDs, No	6		6		0.236
- MDs, Yes	13		4		
TRS	15	6.0 (0 – 38)	3	10.0 (4 – 34)	0.311
DRS	15	0.0 (0 – 6)	3	1.0 (0 – 2)	0.104
Endocrinology (females ≥12 years)	14		5		
Puberty induced, No	7		2		0.620
Puberty induced, Yes	5		3		
- POI, No	2		1		1.000
- POI, Yes	9		3		
Gal-1-P (µmol/g Hb) in erythrocytes*	28	0.36 (0.24 – 0.62)	15	0.40 (0.16 – 0.70)	0.221
Galactitol (mmol/mol creatinine) in urine*	14	128 (109 – 168)	11	175 (113 – 311)	0.204

Notes. Data reported in median and ranges. NBS: newborn screening, IQ: intelligence quotient, MDs: movement disorders, TRS: tremor rating scale, DRS: dystonia rating scale, POI: primary ovarian insufficiency, Gal-1-P: galactose-1-phosphate. * most recent value is presented.

Bone health

In 36 out of 56 patients, the results of a dual-energy X-ray absorptiometry (DEXA) scan were available. In total, 3 out of 17 children (18%) and 2 out of 19 adults (11%) had a bone mineral density ≤ -2 SD at the most recent DEXA scan. The Z-scores of the femoral neck and lumbar spine did not significantly differ between males and females nor between children and adults.

A history of fractures was reported in 7 out of 30 patients. In these patients, the bone mineral density was ≤ -2 SD in two patients and normal in three patients and two patients were too young to undergo a DEXA scan. All reported fractures were preceded by a trauma.

NBS-detected variant patients

A recently identified group of patients present in this cohort are the seven NBS-detected variant patients (**Tables 1** and **2**) of whom four are siblings. Six variant patients have erythrocyte Gal-1-P levels below the detection limit (<0.05 µmol/g Hb) and normal urine galactitol levels on a galactose-restricted diet, while one patient with a residual erythrocyte

GALT activity of 3.6% (just above the limit of quantitation) demonstrated variable Gal-1-P levels ranging from below the detection limit to 0.13 μmol/g Hb as the highest measured level and galactitol levels above the normal range. The most recent Gal-1-P and galactitol levels were significantly lower in the variant patients when compared with the classical patients (p<0.0005). In the four patients with the highest GALT enzyme activities (8.7–9.3%) and normal metabolites on a galactose-restricted diet, the diet was relaxed to a maximum daily allowance of 1200 mg galactose (50 ml dairy product) in three patients and 2400 mg galactose (100 ml dairy product) in one patient. The dietary relaxation did not increase the Gal-1-P or galactitol levels in any of these patients. The assessment of two variant patients was limited due to their young age (4 and 23 months). In the assessed patients, the development of gross and fine motor skills was normal. The speech development was normal in all but one patient, who suffered from persistent adenoiditis for which two adenoidectomies were required in the first two years of life. Currently, none of the variant patients demonstrates long-term complications.

Homozygous p.Ser135Leu patients

In spite of a late diagnosis and late onset of a galactose-restricted diet, both patients do not demonstrate POI or MDs. However, both patients demonstrated a poor intellectual outcome with an IQ well below 85. Both patients have lower Gal-1-P and galactitol levels than classical patients (**Table 1**) and at times Gal-1-P levels even below the detection limit despite dietary incompliance in one patient.

Table 4. Brain MRI abnormalities in patients

Pt ID	Fazekas	Cerebellum	Vermis	Cerebral*	Frontal*	Parietal*	Temporal*	Occipital*	IQ	MD
2	2	+	+	1 – 2	2	2	1	1	77	-
5	1	+	+	1	1	2	0	1	78	-
7	1	-/+	-/+	1	1	1	1	0	78	-
35	1	-	-	0	0	1	0	0	77	Yes
36	0	-	+	1	1	1	1	1	61	-
12	0	+	+	0 – 1	1	1	0	0	81	Yes
14	1	-	+	0	0	0	0	0	83	-
38	1	-	-/+	0 – 1	1	1	0	0	71	-
39	1	+	+	1	1	1	1	1	61	Yes
19	0	-	+	0	0	1	0	0	82	No
20	0	-	+	0	0	1	0	0	88	Yes
24	0	-	-	0	0	0 – 1	0	0	98	No
22	1	-	-	0	0	1	0	0	91	-
23	1	-	-	0	0	0	0	0	93	No
40	0	+	+	0	0	1	0	0	52	Yes
28	0	-	+	0	0	1	0	0	61	No
31	0	-/+	+	0 – 1	1	1	0	0	<50	Yes
42	0	-	-	0 – 1	1	1	0	0	68	Yes

Notes. + atrophy is present, - atrophy is absent, -/+ minor atrophy. * Global Cortical Atrophy (GCA) scale (scores ranging from 0–3 per region), IQ: intelligence quotient, MD: movement disorder. - : data not available, Patient (Pt) ID corresponds with **Table 1**. Patients are listed by decreasing age.

Possible predictors of clinical outcome

Most recent Gal-1-P level

The most recent Gal-1-P level measured by gas chromatography–mass spectrometry was available in 50 patients. There were no significant differences in Gal-1-P levels between children and adults nor between males and females. The Gal-1-P levels between patients with a poor and normal intellectual outcome and with and without MDs and POI were not significantly different. Linear regression analysis indicated that Gal-1-P was not a significant predictor for IQ (p=0.90). Logistic regression to evaluate if Gal-1-P was able to predict POI and MDs was considered unreliable due to broad confidence intervals.

In the screened patient group, patients with a normal intellectual outcome (n=9) had significantly lower Gal-1-P levels than patients with a poor intellectual outcome (n=10) (p=0.017). This group includes the variant patients with higher GALT activities in erythrocytes and significantly lower Gal-1-P levels than the classical patients (p<0.0005). There was a significant negative correlation between the Gal-1-P levels and erythrocyte GALT activity (p<0.0005). In the screened classical patients, there was no significant difference in Gal-1-P levels between patients with a poor and normal intellectual outcome.

In the screened patient group, the Gal-1-P levels between patients with and without MDs and POI demonstrated no significant differences.

Lifetime Gal-1-P

The lifetime Gal-1-P level measured by spectrophotometry was available in 29 patients. The average number of Gal-1-P measurements per patient was 14 (5–59). The lifetime Gal-1-P levels between classical patients with a poor and normal intellectual outcome and with and without MDs and POI were not significantly different.

To investigate whether lifetime Gal-1-P was stable from 12 months, the association with age was tested, demonstrating a significant negative correlation between age and lifetime Gal-1-P (p=0.043). The inclusion of patients with less than five Gal-1-P measurements did not change this finding.

N-glycan profiles

In total, 9 out of 28 glycan peaks (GPs) were significantly different when the samples of adult CG patients were compared with the control samples (p≤0.001). From these nine GPs, a major trend emerges that the bisecting glycans increase in CG patients (GP3, GP21 and GP26, **Figure 1**). Also, an increase in afucosylated bisected glycans (Bn) was observed when compared with controls (**Table 5** and **Supplementary Figure 2A**). There were no significant differences in GPs, G-ratios or IgG *N*-glycan features (core fucosylation and bisected GlcNAcylation) between adult patients with a poor and normal intellectual outcome and patients with or without MDs or POI (data not shown).

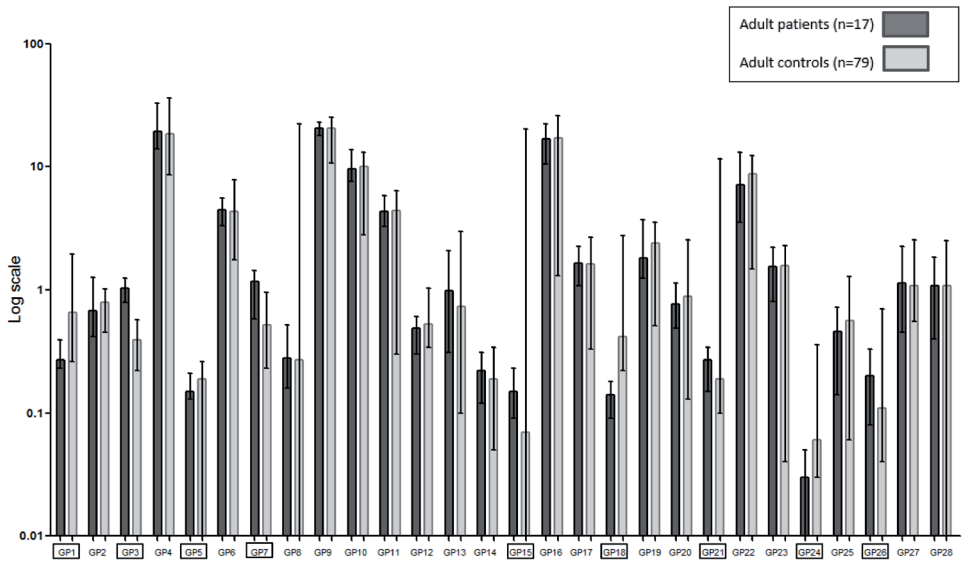


Figure 1. IgG *N*-glycan peaks of adult patients and controls
The percentage areas of 28 IgG *N*-glycan peaks (GPs), quantified in adult galactosemia patients and controls as described in Stockmann et al. 2016. Data reported in median (ranges) with error bars. □: significantly different GPs (p≤0.001), as shown in **Table 5**.

Table 5. *N*-glycans of adult patients

	CG patients (*n*=17)	Controls (*n*=79)	P-value	Main Glycans**
Glycan peaks (GPs)*				
GP1	0.27 (0.23 – 0.39)	0.66 (0.26 – 1.95)	**<0.0005**	FA1
GP3	1.04 (0.79 – 1.25)	0.39 (0.22 – 0.57)	**<0.0005**	A2B
GP5	0.15 (0.13 – 0.21)	0.19 (0.01 – 0.26)	**0.001**	M5
GP7	1.17 (0.58 – 1.44)	0.52 (0.23 – 0.95)	**<0.0005**	A2[3]G1
GP15	0.15 (0.09 – 0.23)	0.07 (0.00 – 20.25)	**<0.0005**	FA2G2
GP18	0.14 (0.09 – 0.18)	0.42 (0.22 – 2.75)	**<0.0005**	FA2G1S1
GP21	0.27 (0.15 – 0.34)	0.19 (0.10 – 11.50)	**<0.0005**	A2BG2S1
GP24	0.03 (0.01 – 0.05)	0.06 (0.03 – 0.36)	**<0.0005**	?
GP26	0.20 (0.08 – 0.33)	0.11 (0.04 – 0.70)	**<0.0005**	A2BG2S2
Glycan features**				
Bn	1.57 (1.19 – 1.69)	0.85 (0.44 – 22.50)	**<0.0005**	Afucosylated neutral glycans

Notes. Data reported in median and ranges. Only significant differences are shown. CG: classical galactosemia, GP: glycan peak, Bn: afucosylated bisected neutral glycans, □: increased in CG patients (when compared with controls), ■: decreased in CG patients (when compared with controls). ? : unknown.
* The percentage areas of 28 IgG *N*-glycan peaks (GPs), quantified as described in Stockmann et al. 2016.
** Main glycans were assigned and *N*-glycan features calculated as described in Pucic et al. 2011.

In the pediatric patients, GP15 (containing core fucosylated biantennary digalactosylated glycan FA2G2 as the major glycan) was significantly decreased and GP21 (containing bisected digalactosylated monosialylated glycan A2BG2S1 as the major glycan) was significantly increased when compared with the nine control samples ($p \leq 0.001$, **Figure 2**, individual data shown in **Supplementary Figure 3A** and **3B**). The core fucosylated bisected neutral glycans (FBn), afucosylated bisected neutral glycans (Bn) and total bisected glycans (B) were significantly higher in pediatric patients when compared with controls (**Table 6** and **Supplementary Figure 2B**). There were no significant differences in GPs or G-ratios between pediatric patients with a poor and normal intellectual outcome and with and without MDs. Both core fucosylated bisected neutral glycans and total bisected glycans were significantly higher in pediatric patients with MDs (**Supplementary Table 1**). The exclusion of the variant patients and homozygous p.Ser135Leu patients did not change any of these results.

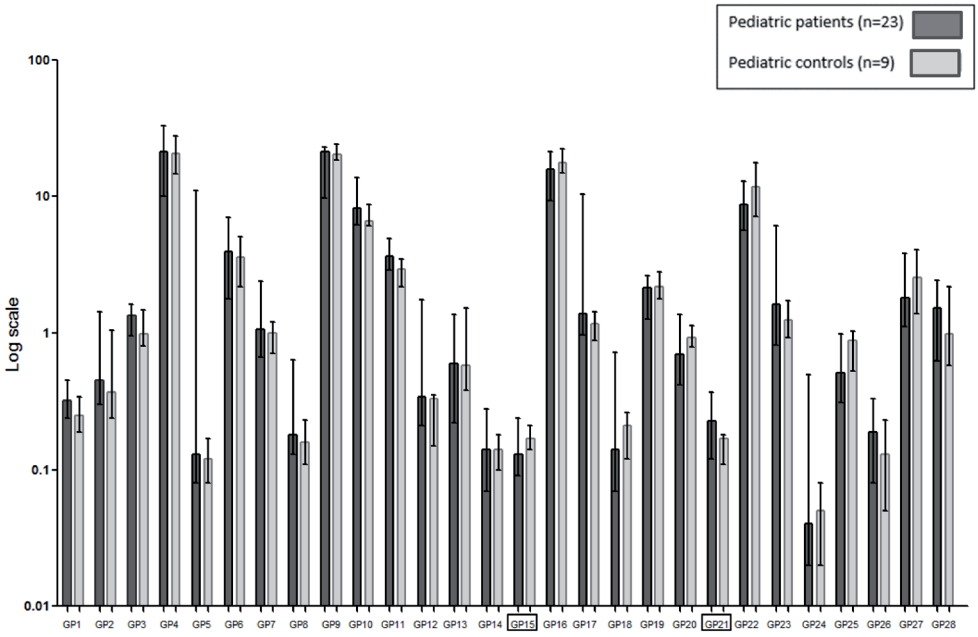


Figure 2. IgG N-glycan peaks of pediatric patients and controls
The percentage areas of 28 IgG *N*-glycan peaks (GPs), quantified in pediatric galactosemia patients and controls as described in Stockmann et al. 2016. Data reported in median (ranges) with error bars.
☐ : significantly different GPs ($p \leq 0.001$), as shown in **Table 6**. Individual data shown in **Supplementary Figure 3A** and **3B**.

Table 6. *N*-glycans of pediatric patients

	CG patients (*n*=23)	Controls (*n*=9)	P-value	Main Glycans**
Glycan peaks (GPs)*				
GP15	0.13 (0.09 – 0.24)	0.17 (0.14 – 0.21)	**0.001**	FA2G2
GP21	0.23 (0.12 – 0.37)	0.17 (0.11 – 0.18)	**< 0.0005**	A2BG2S1
Glycan features**				
FBn	9.55 (7.40 – 18.80)	8.31 (5.72 – 10.25)	**0.022**	Core fucosylated bisected glycans
Bn	1.66 (1.33 – 2.38)	1.25 (1.09 – 1.73)	**0.002**	Afucosylated neutral glycans
B	14.78 (11.35 – 29.49)	12.32 (9.06 – 13.93)	**0.001**	Total bisected glycans

Notes. Data reported in median and ranges. Only significant differences are shown. CG: classical galactosemia, GP: glycan peak, FBn: core fucosylated bisected neutral glycans; FBn: GP6+11+12+17, Bn: afucosylated bisected neutral glycans; Bn: GP3+8+14, B: total bisected glycans; B: GP3+6+8+11+12+14+17+21+23+26+28. ■: increased in CG patients (when compared with controls), □: decreased in CG patients (when compared with controls).

* The percentage areas of 28 IgG *N*-glycan peaks (GPs), quantified as described in Stockmann et al. 2016.

** Main glycans were assigned and *N*-glycan features calculated as described in Pucic et al. 2011.

DISCUSSION

The aim of this cohort study, which includes more than 30% of all Dutch CG patients, was to perform deep phenotyping and to investigate the association between long-term complications and possible predictors of clinical outcome. The results of our study reflect the broad clinical outcome spectrum of CG with an overall below average intelligence and the frequent occurrence of MDs, brain abnormalities on MRI and POI in females. The neurological examination revealed MDs in 47% of patients, which is comparable to the results of Kuiper et al.[33], but less frequent than reported by Rubio-Agusti et al.[34]. This finding might be explained by the fact that the latter reported on adults only, and it confirms the lower frequency of MDs in children, which has also been reported previously[33]. Interestingly, tremors were more frequently and dystonia less frequently observed in our cohort when compared with previous studies[33,34]. Unlike other cohorts, ataxia was not found in our patients.

Patients with a poor intellectual outcome and/or MDs demonstrated more brain abnormalities on MRI than patients without MDs and a higher IQ. The finding that the presence of cerebellar atrophy is associated with IQ is in accordance with previous research that suggested that the cerebellum might be involved in cognition[35,36].

Puberty was induced in 42% of the females aged 12 years and older, and POI was diagnosed in 71%. This frequency is lower when compared with other studies[3,5], possibly because our cohort also includes two homozygous p.Ser135Leu patients without POI.

A low bone mineral density (Z-score ≤ -2 SD) was reported in 14% of the patients, compared with 2.3% in the general population. This finding is in line with previous estimations[37]. We found a higher frequency of low bone mineral density in children than in adults, suggesting later maturation as previously reported in CG[38] and improvement

with age. Considering that no fragility fractures were reported in our cohort, the clinical relevance of the higher incidence of low bone mineral density in CG patients remains uncertain.

Our cohort consists of classical, homozygous p.Ser135Leu and NBS-detected variant patients. In the largest group, the classical patients, the patients with identical genotypes and biochemical outcomes demonstrated highly variable clinical outcomes. The frequency of long-term complications did not differ between screened and non-screened classical patients. Thus, early diagnosis and initiation of treatment does not explain the differences in clinical outcome, which is in line with previous research[4,7,33]. The highly variable clinical outcome spectrum of CG patients and the diversity in genotypes, phenotypes and biochemistry underline the need for predictors of clinical outcome.

In our cohort, Gal-1-P and *N*-glycans as possible predictors of clinical outcome were investigated. Both lifetime and the most recent Gal-1-P levels of patients with and without long-term complications were not significantly different. This seems to be in contrast to the findings of Yuzyuk et al.[14] who evaluated lifetime Gal-1-P. In the screened patients in our cohort, we did find significantly lower Gal-1-P levels in patients with a normal intellectual outcome. However, this resulted directly from the inclusion of the NBS-detected variant patients with a different biochemical profile. In the cohort reported by Yuzyuk et al., the lifetime Gal-1-P levels of patients with a normal long-term outcome were indeed lower. However, some of these patients are comparable to the variant patients in our cohort with higher residual GALT activities and lower Gal-1-P levels. In our cohort, the most recent Gal-1-P levels were able to discriminate between classical and variant patients, but both the lifetime and the most recent Gal-1-P levels were not able to discriminate between classical patients with a poor and normal clinical outcome.

Previous research reported a stabilization in Gal-1-P within a year[14,39]. However, the evaluation of lifetime Gal-1-P in our cohort of classical patients demonstrated a significant negative correlation between lifetime Gal-1-P and age. This indicates a more gradual but further decline in Gal-1-P after the age of 12 months. As it is currently unclear if and when Gal-1-P stabilizes, caution is required for its use as prognostic biomarker.

The evaluation of *N*-glycan profiles demonstrated *N*-glycan variations in CG patients with significant differences between patients and controls for various *N*-glycan peaks. In our cohort, adult patients showed more differentiation between *N*-glycan peaks when compared with controls than pediatric patients. *N*-glycan peak GP21, containing bisecting glycans, as well as afucosylated bisected neutral glycans (Bn), was found to be significantly higher in CG patients when compared with controls in both pediatric and adult patients. This observation conflicts with previous findings[19], where this glycan feature was found to be decreased but does correlate with the more recently published study[20], where *MGAT3* gene expression is significantly upregulated in CG patients.

MGAT3 gene encodes β-1,4-mannosyl-glycoprotein 4-β-*N*-acetylglucosaminyltransferase, which adds bisecting GlcNAcs to *N*-glycans. Possible limitations in all studies are the small sample sizes, and the patients from these studies may have different phenotypes including ethnicity and *GALT* variants. In contrast to previous research, G-ratios were not informative in our cohort of patients[11,17].

In our cohort, both core fucosylated bisected neutral glycans (FBn) and total bisected glycans (B) were significantly higher in pediatric patients with MDs and there may be link to glycan dysregulation in the central nervous system. No other significant differences in *N*-glycans between patients with and without long-term complications were found in our cohort, and it remains unclear if (ongoing) galactosylation abnormalities may be able to predict clinical outcome in CG[17-19]. At this moment, the clinical relevance of the (main) glycan abnormalities found in CG patients is unclear and future research is needed to investigate this further.

Limitations

Even though 56 patients is one of the larger reported CG cohorts, patient numbers are small due to the rarity of the disorder. Furthermore, missing data posed a challenge in the statistical analyses. The use of tremor and dystonia rating scales is hampered because of the lack of control data and age-dependent scales, particularly in the pediatric population where higher scores might represent immature movements due to incomplete brain maturation instead of pathology.

In this study, the fluid-attenuated inversion recovery (FLAIR) sequence was used to assess focal white matter abnormalities on MRI. In certain areas of the brain, most notably deep gray matter and infratentorial structures, T2-weighted images are known to have a higher sensitivity. Unfortunately, this sequence was not available in our MRI protocol.

Even though plasma glycans in individuals have been shown to be stable, a high level of variability has been observed within populations and age-, gender-, environmental- and physiological influences have been reported[40,41]. Moreover, the *N*-glycan profiles consist of multiple variables, which require correction for multiple testing.

Moreover, in this study, both erythrocyte Gal-1-P and serum IgG *N*-glycans were investigated as they are easy to collect and relatively non-invasive when compared with other tissues. The question remains whether Gal-1-P and *N*-glycans in the affected tissues such as brain and ovaries would provide different results. As Gal-1-P is trapped into the cells due to its charged nature, the Gal-1-P levels measured in erythrocytes may not represent the Gal-1-P levels in the affected tissues. Also, serum IgG *N*-glycans may not represent glycosylation patterns in the affected tissues.

The elimination of possible confounders such as age and genetic heterogeneity in the *N*-glycan analyses and genetic heterogeneity in the Gal-1-P analysis reduced the sample size even further and therefore, definitive conclusions on the prognostic value of Gal-1-P and *N*-glycans cannot be drawn.

Strengths

In this study, we included over 30% of the total Dutch CG population. The clinical assessment based on the CG guideline enabled a standardized evaluation, and patients with a second (genetic) diagnosis influencing clinical outcome were excluded. A prospective study design was used for the intelligence testing and neurological examination. Moreover, the neurological examination was performed by one (pediatric) neurologist and complemented with tremor and dystonia rating scales.

Since variant patients and p.Ser135Leu homozygous patients differ from classical CG patients in genotypes, biochemistry and clinical outcomes, which may influence the results, analyses were repeated after the exclusion of the homozygous p.Ser135Leu and variant patients.

Future perspectives

In our cohort, we demonstrated a highly variable clinical outcome with a frequency of long-term complications comparable to previous studies. A long-term complication that has yet received limited attention in the literature is MDs. A standardized neurological evaluation is warranted in all patients, but especially in children with a delay in motor and speech development. Importantly, the impact of the MDs on daily functioning and the possible progressiveness of MDs with age should be investigated.

We hypothesized that the variability in clinical outcome in CG patients is caused by differences in the extent of galactose intoxication and galactosylation abnormalities. However, the results indicate that Gal-1-P and *N*-glycans were not able to predict clinical outcome in our cohort. This may well be the result of limited power and underlines the need for international collaboration to increase patient numbers in studies evaluating predictors of clinical outcome. Importantly, age variability and stability of Gal-1-P should be further investigated.

A remarkable group in our cohort are the variant patients detected since the introduction of NBS with previously unknown genotypes and different clinical and biochemical phenotypes. Currently, all patients with enzyme activities <15% are treated[21]; however the question remains if the variant patients with higher erythrocyte GALT activity (up to 10%) are indeed patients in need of strict dietary treatment, especially since galactose over restriction might be harmful[13]. After the first months of life, the variant patients demonstrated mostly normal Gal-1-P levels even after galactose allowance in some. To determine the optimal treatment of this group, which constitutes 14% of patients detected by NBS in the Netherlands[1], further studies of galactose tolerance are warranted.

In order to develop prognostic biomarkers, improving our understanding of galactose metabolism and the underlying mechanism of long-term complications is crucial. Future studies focusing on galactose metabolism at cell level preferably in the affected tissues and whole body metabolism could provide more insight. In addition, more research into the

brain abnormalities on MRI and the association with clinical outcome may be of value. Visible lesions on MRI are relatively scarce in CG patients and may not satisfyingly explain clinical outcome. The use of diffusion-weighted imaging in CG patients has revealed white matter pathology that correlated with (cognitive) outcome[42]. The use of quantitative MR techniques may contribute to the investigation of pathology in the normal appearing white matter and gray matter and therefore, quantitative analysis of the MR images in this study will be performed.

CONCLUSION

In this study, the deep-phenotyping of a representative cohort of CG patients demonstrated a large variability in clinical outcome. In our cohort, Gal-1-P levels did discriminate between classical and NBS-detected variant patients. However, both Gal-1-P and IgG *N*-glycans were not associated with long-term complications. The variability in clinical outcome necessitates individual and standardized evaluation of all CG patients. Future studies to increase knowledge and understanding of the pathophysiology of CG and its long-term complications are needed to determine the cause of the broad clinical outcome spectrum in CG.

Funding
This study was supported by grants of Stichting Steun Emma, Stofwisselkracht and The Galactosemia Foundation.

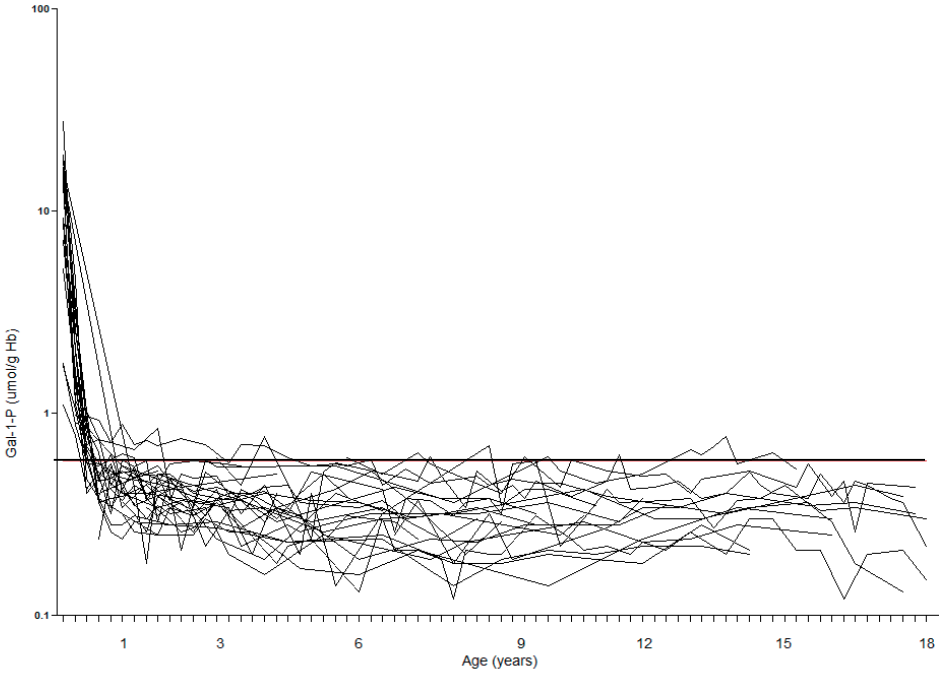
REFERENCES

1. Welling L, Boelen A, Derks TG, Schielen PC, de Vries M, Williams M, et al. Nine years of newborn screening for classical galactosemia in the Netherlands: Effectiveness of screening methods, and identification of patients with previously unreported phenotypes. *Molecular genetics and metabolism*. 2017;120(3):223-8.
2. Bosch AM. Classical galactosaemia revisited. *Journal of inherited metabolic disease*. 2006;29(4):516-25.
3. Waisbren SE, Potter NL, Gordon CM, Green RC, Greenstein P, Gubbels CS, et al. The adult galactosemic phenotype. *Journal of inherited metabolic disease*. 2012;35(2):279-86.
4. Hughes J, Ryan S, Lambert D, Geoghegan O, Clark A, Rogers Y, et al. Outcomes of siblings with classical galactosemia. *Journal of pediatrics*. 2009;154(5):721-6.
5. Coss KP, Doran PP, Owoeye C, Codd MB, Hamid N, Mayne PD, et al. Classical Galactosaemia in Ireland: incidence, complications and outcomes of treatment. *Journal of inherited metabolic disease*. 2013;36(1):21-7.
6. Rubio-Gozalbo ME, Haskovic M, Bosch AM, Burnyte B, Coelho AI, Cassiman D, et al. The natural history of classic galactosemia: lessons from the GalNet registry. *Orphanet journal of rare diseases*. 2019;14(1):86.
7. Fridovich-Keil JL, Walter JH. Part 7: Carbohydrates, Chapter 72: Galactosemia. The Online Metabolic and Molecular Bases of Inherited Disease, OMMBID; Valle D.L., Antonarakis S, Ballabio A, Beaudet A.L., Mitchell G.A.(Eds.). McGraw Hill, New York.
8. Berry GT, Nissim I, Lin Z, Mazur AT, Gibson JB, Segal S. Endogenous synthesis of galactose in normal men and patients with hereditary galactosaemia. *The Lancet*. 1995;346(8982):1073-4.
9. Holton JB. Effects of galactosemia in utero. *European journal of pediatrics*. 1995;154(7 Suppl 2):S77-81.
10. Lai K, Langley SD, Khwaja FW, Schmitt EW, Elsas LJ. GALT deficiency causes UDP-hexose deficit in human galactosemic cells. *Glycobiology*. 2003;13(4):285-94.
11. Coss KP, Byrne JC, Coman DJ, Adamczyk B, Abrahams JL, Saldova R, et al. IgG *N*-glycans as potential biomarkers for determining galactose tolerance in Classical Galactosaemia. *Molecular genetics and metabolism*. 2012;105(2):212-20.
12. Coman DJ, Murray DW, Byrne JC, Rudd PM, Bagaglia PM, Doran PD, et al. Galactosemia, a single gene disorder with epigenetic consequences. *Pediatric research*. 2010;67(3):286-92.
13. Knerr I, Coss KP, Kratzsch J, Crushell E, Clark A, Doran P, et al. Effects of temporary low-dose galactose supplements in children aged 5-12 y with classical galactosemia: a pilot study. *Pediatric research*. 2015;78(3):272-9.
14. Yuzyuk T, Viau K, Andrews A, Pasquali M, Longo N. Biochemical changes and clinical outcomes in 34 patients with classic galactosemia. *Journal of inherited metabolic disease*. 2018;41(2):197-208.
15. Webb AL, Singh RH, Kennedy MJ, Elsas LJ. Verbal dyspraxia and galactosemia. *Pediatric research*. 2003;53(3):396-402.
16. Robertson A, Singh RH, Guerrero NV, Hundley M, Elsas LJ. Outcomes analysis of verbal dyspraxia in classic galactosemia. *Genetics in medicine*. 2000;2(2):142-8.

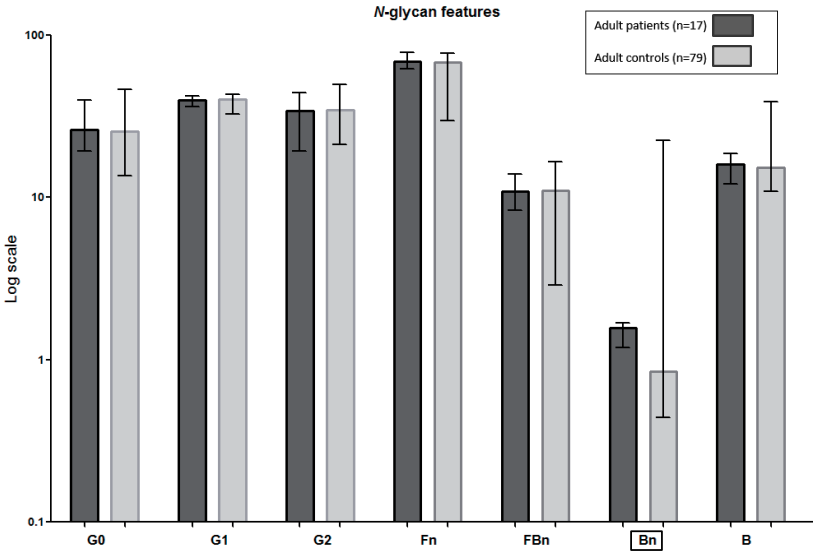
17. Coss KP, Hawkes CP, Adamczyk B, Stockmann H, Crushell E, Saldova R, et al. *N*-glycan abnormalities in children with galactosemia. *Journal of proteome research*. 2014;13(2):385-94.
18. Stockmann H, Coss KP, Rubio-Gozalbo ME, Knerr I, Fitzgibbon M, Maratha A, et al. IgG *N*-Glycosylation Galactose Incorporation Ratios for the Monitoring of Classical Galactosaemia. *JIMD reports*. 2016;27:47-53.
19. Maratha A, Stockmann H, Coss KP, Estela Rubio-Gozalbo M, Knerr I, Fitzgibbon M, et al. Classical galactosaemia: novel insights in IgG *N*-glycosylation and *N*-glycan biosynthesis. *European journal of human genetics*. 2016;24(7):976-84.
20. Colhoun HO, Rubio Gozalbo EM, Bosch AM, Knerr I, Dawson C, Brady J, et al. Fertility in classical galactosaemia, a study of *N*-glycan, hormonal and inflammatory gene interactions. *Orphanet journal of rare diseases*. 2018;13(1):164.
21. Welling L, Bernstein LE, Berry GT, Burlina AB, Eyskens F, Gautschi M, et al. International clinical guideline for the management of classical galactosemia: diagnosis, treatment, and follow-up. *Journal of inherited metabolic disease*. 2017;40(2):171-6.
22. Castor Electronic Data Capture, Ciwit BV, Amsterdam, The Netherlands, 2018.
23. Shin-Bühring Y OM, Ziegler R, Schaub J. A method for galactose-1-phosphate uridyltransferase assay and the separation of its isozymes by DEAE-cellulose column chromatography. *Clinica chimica acta*. 1976;70(3):371-7.
24. Lai K, Langley SD, Singh RH, Dembure PP, Hjelm LN, Elsas LJ, 2nd. A prevalent mutation for galactosemia among black Americans. *Journal of pediatrics*. 1996;128(1):89-95.
25. Welsink-Karssies MM OK, Hermans ME, Hollak CEM, Janssen MCH, Langendonk JG, Oussoren E, Rubio Gozalbo ME, de Vries M, Geurtsen GJ, Bosch AM. Classical Galactosemia: Neuropsychological and psychosocial functioning beyond intellectual abilities. *Orphanet journal of rare diseases*. 2020;15(1):42.
26. Fahn S TE, Marín C. Clinical rating scale for tremor. Parkinson's Disease and Movement Disorders. Baltimore: Williams & Wilkins; 1993.
27. Burke RE FS, Marsden CD, Bressman SB, Moskowitz C, Friedman J. Validity and reliability of a rating scale for the primary torsion dystonias. *Neurology*. 1985;35(1):73-7.
28. Gitzelmann R. Estimation of galactose-I-phosphate in erythrocytes: a rapid and simple enzymatic method. *Clinica chimica acta*. 1969;26(2):313-6.
29. Colhoun HO, Treacy EP, MacMahon M, Rudd PM, Fitzgibbon M, O'Flaherty R, et al. Validation of an automated ultraperformance liquid chromatography IgG *N*-glycan analytical method applicable to classical galactosaemia. *Annals of clinical biochemistry*. 2018;55(5):593-603.
30. Stockmann H, Adamczyk B, Hayes J, Rudd PM. Automated, high-throughput IgG-antibody glycoprofiling platform. *Analytical Chemistry*. 2013;85(18):8841-9.
31. Pucic M, Knezevic A, Vidic J, Adamczyk B, Novokmet M, Polasek O, et al. High throughput isolation and glycosylation analysis of IgG-variability and heritability of the IgG glycome in three isolated human populations. *Molecular and cellular proteomics*. 2011;10(10):M111 010090.
32. Pucic M, Muzinic A, Novokmet M, Skledar M, Pivac N, Lauc G, et al. Changes in plasma and IgG *N*-glycome during childhood and adolescence. *Glycobiology*. 2012;22(7):975-82.

33. Kuiper A, Grunewald S, Murphy E, Coenen MA, Eggink H, Zutt R, et al. Movement disorders and nonmotor neuropsychological symptoms in children and adults with classical galactosemia. *Journal of inherited metabolic disease*. 2019.
34. Rubio-Agusti I, Carecchio M, Bhatia KP, Kojovic M, Parees I, Chandrashekar HS, et al. Movement disorders in adult patients with classical galactosemia. *Movement disorders*. 2013;28(6):804-10.
35. Rapoport M, van Reekum R, Mayberg H. The role of the cerebellum in cognition and behavior: a selective review. *Journal of neuropsychiatry and clinical neurosciences*. 2000;12(2):193-8.
36. Timmann D, Drepper J, Frings M, Maschke M, Richter S, Gerwig M, et al. The human cerebellum contributes to motor, emotional and cognitive associative learning. A review. *Cortex*. 2010;46(7):845-57.
37. van Erven B, Welling L, van Calcar SC, Doulgeraki A, Eyskens F, Gribben J, et al. Bone Health in Classic Galactosemia: Systematic Review and Meta-Analysis. *JIMD reports*. 2017;35:87-96.
38. Waggoner DD, Buist NR, Donnell GN. Long-term prognosis in galactosaemia: results of a survey of 350 cases. *Journal of inherited metabolic disease*. 1990;13(6):802-18.
39. Walter JH, Collins JE, Leonard JV. Recommendations for the management of galactosaemia. UK Galactosaemia Steering Group. *Archives of disease in childhood*. 1999;80(1):93-6.
40. Knezevic A, Polasek O, Gornik O, Rudan I, Campbell H, Hayward C, et al. Variability, heritability and environmental determinants of human plasma *N*-glycome. *Journal of proteome research*. 2009;8(2):694-701.
41. Gornik O, Wagner J, Pucic M, Knezevic A, Redzic I, Lauc G. Stability of *N*-glycan profiles in human plasma. *Glycobiology*. 2009;19(12):1547-53.
42. Timmers I, Zhang H, Bastiani M, Jansma BM, Roebroeck A, Rubio-Gozalbo ME. White matter microstructure pathology in classic galactosemia revealed by neurite orientation dispersion and density imaging. *Journal of inherited metabolic disease*. 2015;38(2):295-304.

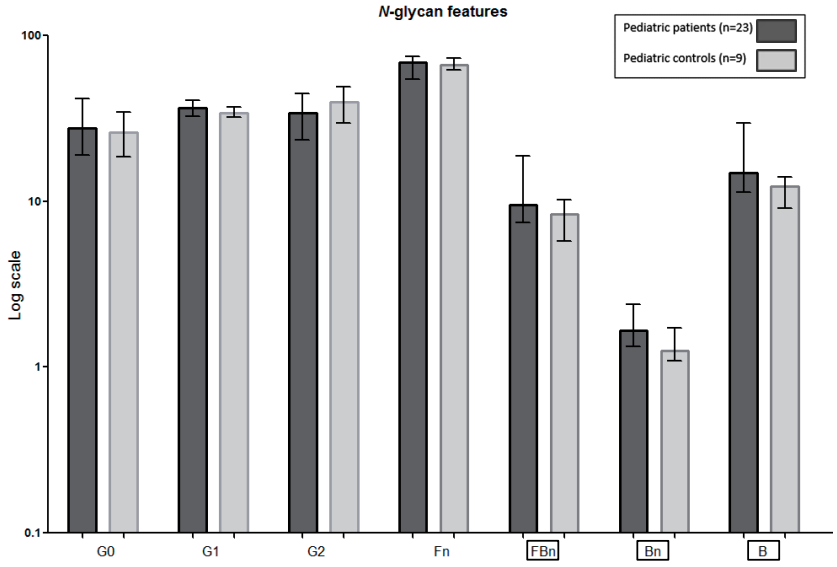
SUPPLEMENTARY MATERIAL



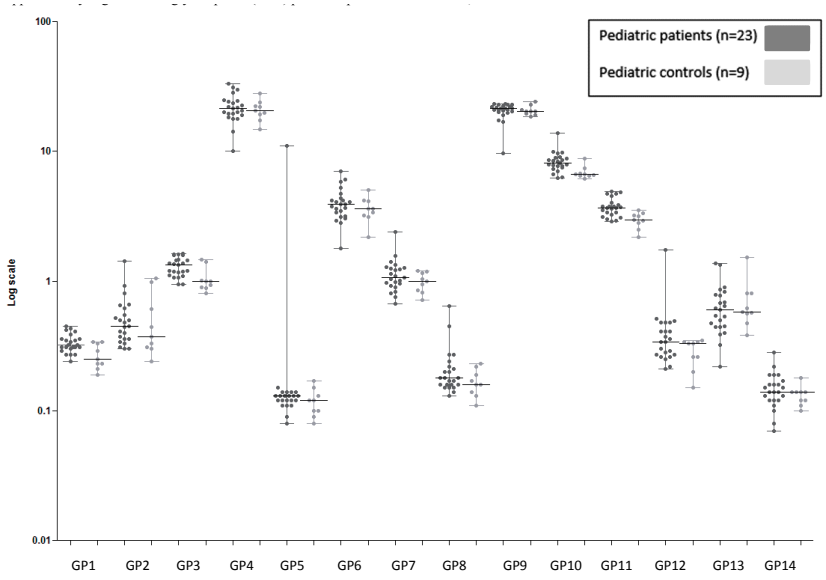
Supplementary Figure 1. Gal-1-P levels of classical patients
Galactose-1-phosphate (Gal-1-P) measurements in erythrocytes of classical patients from birth to adulthood, demonstrated on a log scale. The Gal-1-P levels measured by spectrophotometry are below 0.58 µmol/g Hb in diet adherent patients, which is indicated by the horizontal line.



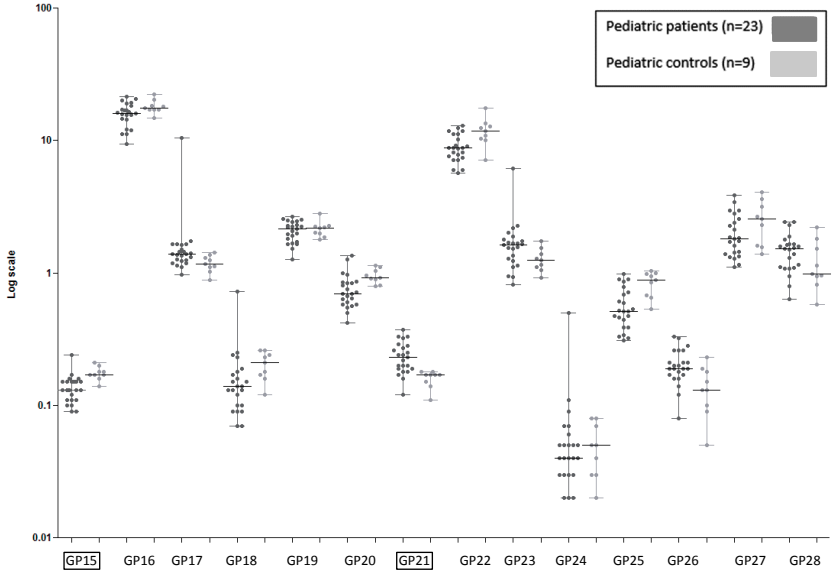
Supplementary Figure 2A. *N*-glycan features of adult patients and controls
N-glycan features were calculated as described in Pucic et al. 2011. G0: agalactosylated structures, G1: monogalactosylated structures, G2: digalactosylated structures, FBn: core fucosylated bisected neutral glycans; FBn: GP6+11+12+17, Bn: afucosylated bisected neutral glycans; Bn: GP3+8+14, B: total bisected glycans; B: GP3+6+8+11+12+14+17+21+23+26+28, □: significantly different *N*-glycan features, as shown in **Table 5**.



Supplementary Figure 2B. *N*-glycan features of pediatric patients and controls
N-glycan features were calculated as described in Pucic et al. 2011. G0: agalactosylated structures, G1: monogalactosylated structures, G2: digalactosylated structures, FBn: core fucosylated bisected neutral glycans; FBn: GP6+11+12+17, Bn: afucosylated bisected neutral glycans; Bn: GP3+8+14, B: total bisected glycans; B: GP3+6+8+11+12+14+17+21+23+26+28, □: significantly different *N*-glycan features, as shown in **Table 6**.



Supplementary Figure 3A. *N*-glycan peaks (1-14) of pediatric patients and controls, individual data
The percentage areas of IgG *N*-glycan peaks (GPs) 1-14, quantified in pediatric galactosemia patients and controls as described in Stockmann et al. 2016. Data reported in median (ranges) with error bars.
□: significantly different GPs (p≤0.001), as shown in **Table 6**.



Supplementary Figure 3B. *N*-glycan peaks (15-28) of pediatric patients and controls, individual data
The percentage areas of IgG *N*-glycan peaks (GPs) 15-28, quantified in pediatric galactosemia patients and controls as described in Stockmann et al. 2016. Data reported in median (ranges) with error bars.
□: significantly different GPs (p≤0.001), as shown in **Table 6**.

Supplementary Table 1. *N*-glycan features of pediatric patients with and without MD's

N-glycan Feature*	MD (*n*=8)	No MD (*n*=12)	P-value
FBn	10.66 (8.49 – 13.64)	8.82 (7.55 – 18.80)	**0.045**
B	15.36 (14.22 – 18.91)	14.00 (11.35 – 29.49)	**0.037**

Notes. Data reported in median and ranges. MD: movement disorder, FBn: core fucosylated bisected neutral glycans, B: total bisected glycans, ▭: increased in patients with a movement disorder.
* *N*-glycan features were calculated as described in Pucic et al. 2011.

Chapter 3

Cognitive functioning in patients with classical galactosemia: a systematic review

M.E. Hermans, M.M. Welsink-Karssies, A.M. Bosch,
K.J. Oostrom, G.J. Geurtsen

Orphanet Journal of Rare Diseases. 2019 Oct;14(1):226.

ABSTRACT

Background

Patients with the metabolic disorder classical galactosemia suffer from long-term complications despite a galactose-restricted diet, including a below average intelligence level. The aim of the current review was to investigate the incidence and profile of cognitive impairments in patients with classical galactosemia.

Methods

MEDLINE, EMBASE and PsychINFO were searched up to 23 October 2018 for studies examining information processing speed, attention, memory, language, visuospatial functioning, executive functioning and social cognition in patients with confirmed classical galactosemia utilizing standardized neuropsychological tests. Data synthesis followed a narrative approach, since the planned meta-analysis was not possible due to large variability between the neuropsychological assessments.

Results

Eleven studies were included, including case-studies. The quality of most studies was moderate to low. As a group, patients with classical galactosemia exhibit below average to low scores on all cognitive domains. A large proportion of the patients perform on an impaired level on attention, memory and vocabulary. Evidence for impairments in information processing speed, language, visuospatial functioning and aspects of executive functioning was limited due to the small number of studies investigating these cognitive functions. Social cognition was not examined at all.

Conclusion

Given the moderate to low quality of the included studies and the limited evidence in many cognitive domains, the incidence of cognitive impairment in patients with classical galactosemia is not yet clear. Both clinicians and researchers encountering patients with classical galactosemia need to be aware of possible cognitive impairments. Future well-designed studies are needed to determine the cognitive profile of classical galactosemia. This can be the basis for the development of intervention strategies.

INTRODUCTION

Classical Galactosemia (CG; OMIM: 230400) is a rare autosomal recessive metabolic disorder with an incidence between 1:16.000 and 1:60.000 in Europe and the USA[1]. The disorder is caused by a deficiency of galactose-1phosphate uridylyltransferase (GALT, EC 2.7.7.12). Due to this deficiency, newborns develop a life-threatening illness after the ingestion of breast milk or formula. If CG is suspected, a galactose-restricted diet is started which eliminates the acute clinical symptoms[2]. CG is confirmed by means of absent or barely detectable GALT enzyme activity in red blood cells and/or the presence of two pathogenic *GALT* mutations[3]. Despite the lifelong diet, patients with CG develop long-term complications in different degrees of severity. Besides neurological deficits and primary ovarian insufficiency[1], cognitive functioning seems to be affected. A recent meta-analysis of intellectual functioning in a sample of early-treated patients with a confirmed diagnosis of CG found a mean below average total intelligence quotient (IQ) of 87. Only 15% had an average or above average IQ (100 or higher) in contradiction to the 50% found in the general population[4]. The lower level of cognitive functioning seems to affect the health-related quality of life[5] and a large proportion of the patients need additional care and guidance in the domain of mental functions[6]. Unfortunately, the majority of articles investigating the cognitive complications of CG only used intelligence tests which are designed to assess overall cognitive functioning[7] and not specific cognitive domains (*i.e.* information processing speed, executive functioning, memory, language, visuospatial functioning and social cognition). A low IQ can be caused by a global impairment in the general mental abilities, but it can also be the result of a specific impairment in one or multiple cognitive domains lowering scores of some or several subtests and consequently the IQ[8]. Consequently, it is important to delineate the cognitive profile of CG to understand the process underlying the lower level of cognitive functioning, to improve prognostic accuracy and to identify cognitive areas in which additional guidance and/or rehabilitative interventions are needed. Therefore, the aim of the current review is to systematically investigate cognitive functioning in patients with CG in order to answer the following questions:

1. What is the incidence of cognitive impairment in classical galactosemia?
2. Which cognitive domains are impaired in patients with classical galactosemia?

METHODS

The current systematic review was performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) method[9]. The PRISMA checklist of this review can be found in **Supplementary Table 1**.

Search strategy

The electronic databases MEDLINE, EMBASE and PsychINFO were searched up to 23 October 2018 with a medical information specialist. The search strategies for MEDLINE and EMBASE were developed to target the patient population and were modified by manually omitting irrelevant clusters of related articles identified by VOSviewer (see **Supplementary Figure 1**[10]). The final search strategies are included in **Supplementary Table 2**. In addition, reference lists of included articles and (systematic) reviews were hand searched. All records were de-duplicated in EndNote and all unique results were uploaded to the systematic review software program Covidence.

Eligibility criteria

Both the title- and abstract screening and the subsequent full-text screening were independently performed by MEH and MMWK. Disagreement was resolved by consensus and consultation of GJG or AMB. Studies were included if they investigated patients with CG, confirmed by either genetic analysis with two pathogenic mutations and/or absent or barely detectable red blood cell GALT enzyme activity. Studies that selected a specific subgroup of CG patients based on clinical outcome were excluded. Moreover, studies needed to report standardized results of standardized neuropsychological tests examining cognitive domains. A standardized neuropsychological test requires standardized administration and scoring procedures, and the presence of normative data[8]. Studies solely assessing general cognitive status or intelligence were excluded. Studies only reporting aggregated scores of test batteries were also excluded.

Given the expected relatively small number of studies, there were no restrictions of age. Full-text, original articles of any publication year written in either English or Dutch were included. If multiple studies reported on the same patient cohort, the study reporting on the largest proportion of the cohort was included.

Data extraction

Data on study characteristics (*i.e.* sample size, study design, control characteristics), patient characteristics (*i.e.* age, gender, criteria for diagnosis, age of diagnosis, age of start diet, compliance to diet, genetic mutation, clinical outcome and psychiatric symptoms) and cognitive outcomes were extracted from the included studies by both MEH and MMWK independently. Age of diagnosis and start of diet were included in the data extraction since late initiation of the galactose-restricted diet (*i.e.* after eight weeks) has been found to be related to lower intelligence levels in CG patients[11,12]. Moreover, the specific genetic mutations reported in the studies were extracted since some pathogenic mutations (*i.e.* homozygous p.Ser135Leu) are associated with a less severe clinical outcome[13]. Clinical outcome was extracted since several outcomes might influence the performance on neuropsychological tests, including severe mental retardation, speech disorders and motor impairments[8] as well as psychiatric symptoms, especially mood

disorders[14].

Four authors were contacted for numerical data that was solely described in their papers. One author responded within the set period of 1.5 months[15]. All neuropsychological (sub-)tests were categorized into different cognitive domains (*i.e.* information processing speed, attention, memory, language, visuoconstruction and visuospatial functioning, executive functioning and social cognition), based on neuropsychological handbooks[8,16]. In the current review, test scores below the ninth percentile (*i.e.* z-score ≤ -1.4) are described as impaired. A cognitive domain is defined as impaired if the results of multiple tests categorized in that particular cognitive domain are below the ninth percentile. For longitudinal designs, only the data of the first measurement were included.

Risk of bias assessment

Risk of bias assessment was performed by MEH and verified by GJG. The Scottish Intercollegiate Guidelines Network (SIGN) quality appraisal checklists[17] were used for case-control studies. The critical appraisal checklists of the Joanna Briggs Institute (JBI[18]) were used for the remaining study designs. The assessment was done while taking the low incidence of the disease into account.

Data synthesis

The outcome measures and quality of report of the results of all included studies were evaluated first. If multiple articles examined the same cognitive function with relatively similar tests and the quality of the studies was sufficient, a meta-analytic approach was planned by means of a random-effects model because of the expected heterogeneity between studies. If a quantitative approach was not feasible because of above-mentioned reasons, a narrative approach was planned.

RESULTS

Study selection

The database-search yielded a total of 6142 records. After the removal of duplicates, the remaining 4144 records were screened. Hand-searching the reference lists resulted in no extra articles. Finally, eleven studies were retained (see **Figure 1**), including two case-control studies[19,20], five cross-sectional studies[15,21-24], one case-series[25] and three case reports[26-28]. Several studies investigating the cognitive development of patients with CG were excluded, since they only used developmental scales and/or intelligence tests, which are not informative about the level of functioning in a specific cognitive domain.

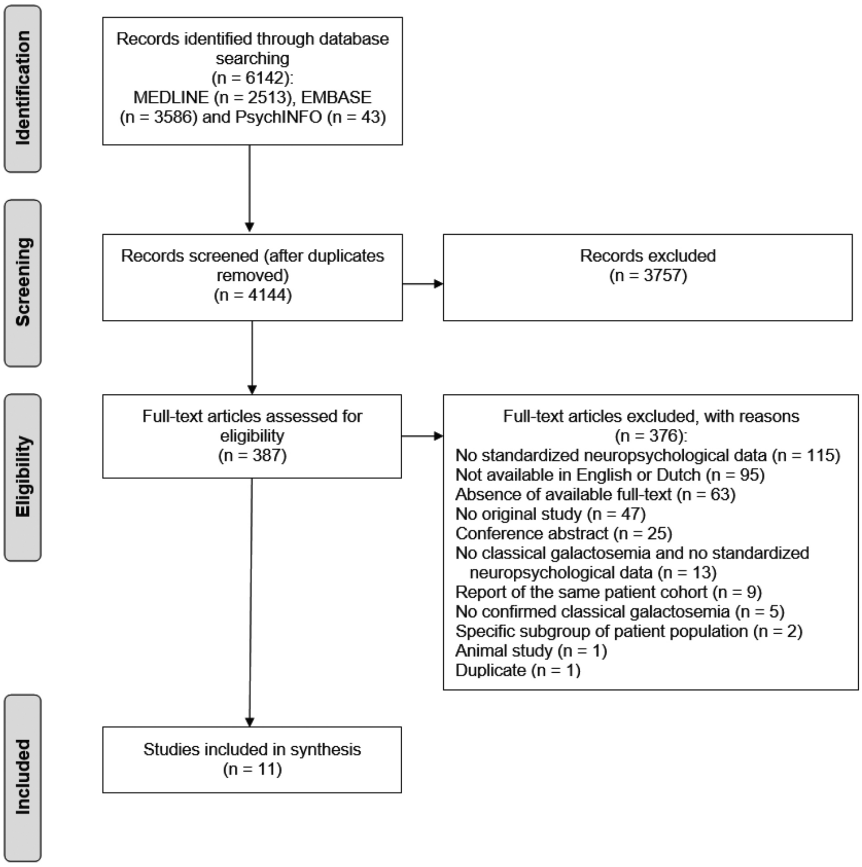


Figure 1. PRISMA flow diagram of the study selection process of the current systematic review regarding cognitive functioning in patients with CG[9].

Study characteristics

The sample consisted of 177 patients with CG (see **Table 1**). The study sample sizes varied from one to 45. Six studies included only children, and five studies included both children and adults with an age range between two and 53 years. The genetic mutation was known of 71 patients, of which 48 were homozygous for the Q188R mutation, which might be related to a more severe outcome[1]. Due to the inclusion of four studies that did not perform or report the results of genetic analysis[21-23,26], the presence of patients carrying the homozygous p.Ser135Leu variant remained unknown. Two studies reported the presence of movement disorders, in which tremor and ataxia were the most common symptoms[15,23]. Psychiatric symptoms were reported in two studies[15,19] and lower intellectual functioning in the majority of the studies. All studies used normative data to evaluate cognitive functioning. Two studies additionally used control subjects[19,20]. Both of these studies matched their controls on age and gender, and one of them added parental educational

coding as a matching variable[20]. Of one study, individual patient data were available making it possible to exclude six individual patients that received the diagnosis CG after 56 days (*i.e.* 8 weeks) in order to avoid any influence of late treatment[15].

Risk of bias assessment

According to the JBI- and the SIGN checklists, only one study was found to be of high quality[19]. Three studies were of low quality, with a high risk of bias[20,26,28]. All other studies were found to be of moderate quality. The results of the risk of bias assessment can be found in **Supplementary Table 3**. The most frequent issue was the recruitment process of the patients in the eight studies investigating multiple patients. Two studies had a high risk for selection bias due to nonconsecutive- and incomplete inclusion[25], or unclear exclusion criteria for controls and cases[20]. In all five cross-sectional studies the recruitment process was unclear[15,21-24]. Only one study described the recruitment process in sufficient detail[19]. Moreover, three studies applied exclusion criteria containing neurological or psychiatric disorders[21] and mental retardation[19,22]. Another major issue was the scarce report of the patients' age at the initiation of the galactose-restricted diet, making it impossible to infer whether the results might be influenced by late treatment. The presence of psychiatric symptoms was mentioned in three studies but possible effects for, or associations with cognitive outcomes were not tested[15,19,21]. Moreover, in the majority of the included studies the association between IQ and performance on the neuropsychological tests was not tested nor accounted for. Finally, there was a large variability between studies in utilized neuropsychological tests. Based on neuropsychological handbooks[8,16], four studies used tests of moderate psychometric quality and/or with older normative data[15,21,23,26]. One study used a test of unclear psychometric quality[22]. Due to this variability and the moderate to low quality of the included studies, a quantitative meta-analytic approach was not possible. A systematic, narrative approach was utilized for the current review.

Cognitive outcomes

All results of the neuropsychological tests can be found in **Table 2**.

Information processing speed

Two studies examined information processing speed[21,23]. A cross-sectional study reporting a total score of two subtests measuring information processing speed and cognitive inhibition, demonstrated an impaired performance averaged across a sample of 24 patients (adults and children), but did not report what process caused the impairment[21]. Thirty percent of the individual patients performed on an impaired level (*i.e.* 29.1%) in contradiction to about 8% in the normal population[16]. Another cross-sectional study found an impaired visual information processing speed in children (z= -1.86) and a below average result in adults (z= -1.33)[23].

Table 1. Study characteristics of the included studies

Author (year)	Study design	*n* [P]/[C]	Mean age y:mo (range)	Gender	Criteria for diagnosis	Age of diagnosis/ start of diet (n)	Genetic mutation (n)	Clinical outcome (n)	Psychiatric symptoms (n)
Antshel et al. (2004)[19]	Case Control	25 / 20	10:9 (8-14)	10 females, 15 males	Confirmed by genetic analysis	NR / first week of life	Q188R/ Q188R (25)	Special education (14), IQ: 84.3 (SD=8.3)	CDI: z= 0.14 (SD=0.04) ADHD (4): no medication 24h before NPA
Doyle et al. (2010)[21]	Cross-sectional	28	29:3 (15-53)	20 females, 8 males	Confirmed by very low GALT-activity in leucocytes	Neonatally / NR[a]	NI	IQ: 88.9 (SD=16.8)	No psychiatric disorders.
Hoffmann et al. (2011)[22]	Cross-sectional	32	21:2 (9:9-37:4)	12 females, 20 males	Confirmed by GALT activity <2% and genetic analysis	NR / NR	NR	IQ: 76.2 (SD=14.8)	NR
Iakovou et al. (2018)[26]	Case Report	1	25	1 male	Confirmed by absent GALT activity	Neonatally/ Neonatally	NI	IQ: 115 (age of 19)	NR
Kaufman et al. (1995)[23]	Cross-sectional	45	4:0-39	22 females, 23 males	Confirmed by absence of GALT activity	Neonatally or in infancy / Neonatally or in infancy	NI	Tremor, ataxia and dysmetria (12)	NR
Lewis et al. (2012)[25]	Case series	4	7:7 7:8 8:5 9:2	3 females, 1 male	Confirmed by GALT activity < 1.3 units/ml blood and genetic analysis	5 days / 5 days (3) or birth (1)	Q188R/ Q188R (4)	IQ: 118 IQ: 67 IQ: 60 IQ: 73	NR
Lewis et al. (2013)[20]	Case Control	1 / 3	5:7	1 male	Confirmed by genetic analysis	6 days / 6 days	Q188R/ Q188R	IQ: within average range	NR

Lewis et al. (2013)[27]	Case Report	1	2:1	1 female	Confirmed by GALT activity <0.01U/g Hb and genetic analysis	7 days / 7 days	Q188R/ Q188R	Learning disabled IQ=60	NR
Ng et al. (2003)[28]	Case Report	1	8	1 female	Confirmed by GALT activity (0.1-0.3%) and genetic analysis	± 2-3 weeks / ± 2-3 weeks	Q133R/ other	IQ: 123 Normal ovarian function later in life	NR
Van Erven et al. (2017)[24]	Cross-sectional	12	17:4 (14:6-21:1)	9 females, 3 males	Confirmed by GALT activity (0.55%) and/ or genetic analysis.	NR / Mean of 11 days (range: 0-35)	Q188R/ Q188R (5) Q188R/ other (1) L195P/ K229N (3) W134fs/ unknown (2) Unknown (1)	Special education (9)	NR
Waisbren et al. (2012)[15]	Cross-sectional	27	29:10 (18-51)	13 females, 14 males	Confirmed by absence of GALT activity and genetic analysis.	Mean of 12 days (range: 1-42)/NR	Q188R/ Q188R (12) Q188R/ other (12) K285N/ L195P (1) K127Q/ deletion (1) Deletion/ deletion (1)	IQ: 89.4 (SD=19.5) Tremor (11) Ataxia (3) Dysarthria (6) Apraxia (2)	Depression (10) Anxiety (20)

Notes. n: sample size, p: patient group, c: control group, NBS: newborn screening, NR: not reported, IQ: total intelligence quotient, SD: standard deviation, CDI: children depression inventory, z: z-score, ADHD: attention deficit hyperactivity disorder, NPA: neuropsychological assessment, NI: not investigated.

[a] The diet of one patient was relaxed from the age of three years old onward.

Attention

Two cross-sectional studies addressed attention and found no impairment[21,23]. However, in both studies the range of performances exceeded the level of impairment, indicating that a proportion of the patients (*i.e.* 29.6%)[21] performed on an impaired level.

Memory

Two studies addressed verbal memory[21,23]. A cross-sectional study of children and adults found no impairments in verbal information encoding and retrieval[21]. However, 32.1% of the patients performed on an impaired level on encoding and 21.4% on retrieval. The same was found in another cross-sectional study[23].

Three studies examined visual memory[19,21,24]. A pediatric case-control study found no impairments on both the immediate- and delayed recall of the structural elements of a complex figure, but the immediate- and delayed recall of incidental elements of the figure was significantly lower in the patient group than in the control group (p<0.001) and impaired (z= -1.47 and z= -1.43)[19]. A small cross-sectional study reported an impaired overall immediate recall of the same complex figure but did not differentiated the results of the different elements or reported the delayed recall-results[24]. Although no memory impairment was found in 28 adults and children in another cross-sectional study, approximately 40% of the patients scored on an impaired level on both immediate- and delayed recall[21].

Language

Six studies examined expressive language[19,20,22,23,25,27]. No impairment in expressive vocabulary was found in a case-control sample of children although the patient group differed significantly from the control group (p<0.001). No difference was present if a phonemic cue was presented (p=0.227)[19]. A cross-sectional study found expressive vocabulary impairment in adults, but not in children[23]. However, the range of performances was large indicating that a proportion of the patients functioned on an impaired level. A case-series reported expressive vocabulary impairment in two out of four children[25]. A poorly designed case-control study found no impairment in another five-year old patient[20].

One cross-sectional study of 32 children and adults found impairments on another aspect of expressive language, namely repetition, measured by a German test of unclear psychometric quality[22]. Lastly, two case studies assessed multiple aspects of expressive language by means of a language scale[20,27]. Both found an impairment in expressive language.

Table 2. Results of neuropsychological assessment reported in the included studies

Author (year)	Tests	Memory	Information Processing Speed	Attention/Executive functioning	Language	Visuospatial functioning & construction
Antshel et al. (2004)[19]	ROCF BNT WCST BVMI	ROCF IR organization: not impaired (z= -0.57 (SD=0.19); p=0.307) ROCF IR structural elements accuracy: not impaired (z= -0.25 (SD=0.11); p=0.338) ROCF IR incidental elements accuracy: impaired (z= -1.47 (SD=0.53); p<0.001) ROCF DR organization: not impaired (z= -0.58 (SD=0.17); p=0.387) ROCF DR structural elements accuracy: not impaired (z= -0.63 (SD=0.08); p=0.316) ROCF DR incidental elements accuracy: impaired (z= -1.43 (SD=0.32); p<0.001)	NI	WCST completed categories: not impaired (z= -1.32 (SD=0.43); p<0.001) WCST perseverative errors (%): impaired (z= -1.98 (SD=0.38); p<0.001)	BNT total correct: not impaired (z= -0.89 (SD=0.18); p <0.001) BNT total correct with phonemic cue: not impaired (z= 0.43 (SD= 0.16); p=0.227)	ROCF Copy organization: not impaired (z= -0.18 (SD= 0.10); p=0.421) ROCF Copy structural elements accuracy: not impaired (z= -0.26 (SD=0.04); p=0.309) ROCF Copy incidental elements accuracy: not impaired (z= -1.35 (SD=0.45); p<0.001) BVMI: not impaired (z= -0.77 (SD=0.27); p=0.089)
Doyle et al. (2010)[21]	Hayling Test BSAT WMS III VOSP - Incomplete letters - Position discrimination	WMS III Auditory IR Index: not impaired (z= -0.53 (SD=1.07), 32.1% < P9) WMS III Auditory DR Index: not impaired (z= -0.37 (SD=1.13), 21.4% < P9) WMS III Auditory Recognition Index: not impaired (z= -0.25 (SD=0.91), 7.1% < P9) WMS III Visual IR Index: not impaired (z= -1.27 (SD=0.73), 39.4% < P9) WMS III Visual DR Index: not impaired (z= -1.18 (SD=0.88), 39.3% < P9)	Hayling Test (n=24): impaired (z= -1.83 (SD= 0.4), 29.1% < P9)	WMS III Digit Span (n=27): not impaired (z= -1.03 (SD=0.73), 29.6% < P9) WMS III Spatial Span (n=25): not impaired (z= -0.93 (SD=1.07), 28.0% < P9) WMS III Letter-number sequencing (n=27): not impaired (z= -1.07 (SD=0.87), 29.7% < P9) BSAT (n=24): impaired (z= -1.47 (SD=0.37), 16.7% < P9) Hayling Test (n=24): impaired (z= -1.83 (SD=0.4), 29.1% < P9)	NI	VOSP Incomplete letters (n=23): not impaired (z= 0 (SD=1.0), 4.3% <P9) VOSP Position discrimination (n=23): impaired (z= -2.13 (SD=3.38), 43.4% <P9)

Hoffmann et al. (2011)[22]	HWWRT	NI	NI	NI	Impaired: 84.4% of the patients had errors (> cut-off of 1)	NI
Iakovou et al. (2018)[26]	PPVT-III	NI	NI	NI	PPVT-III: not impaired (z= 1.20)	NI
Kaufman et al. (1995)[23]	WJ-R COG PPVT-R BVMI	WJ-R Memory for names (n=40): Children: not impaired (z= -0.79 (SD=1.09)) Adults: not impaired (z= -1.33 (SD=0.91))	WJ-R Visual Matching (n=40): Children: impaired (z= -1.86 (SD=1.84)), Adults: not impaired (z= -1.33 (SD=1.27))	WJ-R Analysis-synthesis (n=40): Children: not impaired (z= -1.18 (SD=0.97)), Adults: not impaired (z= -0.99 (SD=1.01)) WJ-R Memory for sentences (n=40): Children: not impaired (z= -1.19 (SD=1.29)), Adults: not impaired (z= -1.21 (SD=0.93))	WJ-R Incomplete words (n=40): Children: not impaired (z= -1.03 (SD=0.96)), Adults: not impaired (z= -0.69 (SD=1.03)) WJ-R Picture vocabulary (n=40): Children: not impaired (z= -1.24 (SD=1.06)), Adults: impaired (z= -1.81 (SD=1.24)) PPVT-R (n=36): Children: not impaired (z= -0.93 (SD=0.89)), Adults: impaired (z= -1.57 (SD=1.09))	WJ-R Visual closure (n=40): Children: not impaired (z= -0.27 (SD=1.0)), Adults: not impaired (z= -0.84 (SD=0.85)) BVMI (n=36): Children: not impaired (z= -1.25 (SD=0.87))
Lewis et al. (2012)[25]	PPVT-4 EVT-2	NI	NI	NI	PPVT-4: Child 2,3,4 impaired (z= -1.67; z= -2.2; z= -2.0), Child 1 not impaired (z= 0.73) EVT-2: Child 3,4 impaired (z= -1.93; z= -2.0), Child 1,2 not impaired (z= 0.47; z= -1.33)	NI

Lewis et al. (2013)[20]	PLS-4 PPVT-4 EVT-2	NI	NI	NI	PLS-4 Auditory comprehension: not impaired (z= -1.13; p≥0.05) PLS-4 Expressive communication: impaired (z= -1.67; p ≥ .05) PPVT-4: not impaired (z= -0.47; p≥.05) EVT-2: not impaired (z= -1.07; p≥0.05)	NI
Lewis et al. (2013)[27]	PLS-4	NI	NI	NI	PLS-4 Receptive communication: impaired (z= -2.27) PLS-4 Expressive communication: impaired (z= -1.67)	NI
Ng et al. (2003)[28]	PPVT-R	NI	NI	NR	PPVT-R: not impaired (z= 1.13)	NR
Van Erven et al. (2017)[24]	ROCF	ROCF IR: impaired (z= -2.0)	NI	NI	NI	NI
Waisbren et al. (2012)[15]	PPVT-4	NI	NI	NI	PPVT-4: not impaired (z= -0.87 (SD=0.98))	NI

Notes. IR: immediate recall, DR: delayed recall, z: z-score, *SD*: standard deviation, p: p-value, NI: not investigated, NR: not reported, P2: Second percentile, ROCF: Rey-Osterrieth Complex Figure Test, BNT: Boston Naming Test, WCST: Wisconsin Card Sorting Test, BVMI: Beery Developmental test of Visual-Motor Integration, WRAML: Wide Range Assessment of Memory and Learning, BSAT: Brixton Spatial Anticipation Test, WMS III: Wechsler Memory Scale III, VOSP: Visual Object and Space Perception Battery, HWWRT: Hierarchische Wortlisten word-repetition test, PPVT-III: Peabody Picture Vocabulary Test-3, WJ-R COG: Woodcock-Johnson Psychoeducational Battery - Revised: Tests of cognitive ability, PPVT-R: Peabody Picture Vocabulary Test - Revised, PPVT-4: Peabody Picture Vocabulary Test-4, EVT-2: Expressive Vocabulary Test-2, PLS-4: Preschool Language Scale-4.

Seven studies addressed receptive language[15,20,23,25-28]. No receptive vocabulary impairment was found in a cross-sectional study concerning 27 early-treated adults[15]. In contrast, another cross-sectional study concerning both adults and children found impairment in adults, but not in children[23]. However, the scores of patients in both studies showed a large variation indicating that a proportion of the patients performed on an impaired level. A case-series found receptive vocabulary impairment in three children and no impairment in one child[25]. Three case reports, of which one was controlled, found no impairment[20,26,28]. A group study reported no impairments in phonological awareness, another basic aspect of receptive language[23]. Lastly, two case studies assessed multiple aspects of receptive language by means of a language scale[20,27]. Both found an impairment in receptive language.

Visuospatial functioning

Two studies addressed visuoconstruction[19,23]. A case-control study found no impairment in copying a complex figure, but the patients performed significantly worse on copying the incidental parts of the figure in comparison to the controls (p<0.001)[19]. They also showed no impairment on another copying test, a result also found in another cross-sectional study[23].

Two studies addressed visual perception[21,23]. Averaged across all 23 patients, impairment in space perception, but not in object perception was found in a cross-sectional study[21]. Only 4.3% of the patients performed on an impaired level on object perception in contradiction to the 43.4% on space perception. Another cross-sectional study found no impairment in object perception, but scores differed extensively between individual patients[23].

Executive functioning

Three studies examined executive functioning[19,21,23], which is an umbrella term for several higher-order functions of which four were investigated in patients with CG (*i.e.* working memory, abstract thinking, cognitive flexibility and cognitive inhibition). A cross-sectional study found no working memory impairment after averaging across all patients, but 30% of the patients performed on an impaired level[21]. A pediatric case-control study assessed abstract thinking together with cognitive flexibility[19]. The patients performed worse than controls on both abstract thinking and cognitive flexibility (p<0.001), but only the performance on cognitive flexibility reached a level of impairment (z= -1.98) indicating impairment in cognitive flexibility alone. Another cross-sectional study also utilized a test measuring both abstract thinking and cognitive flexibility[21]. They found an impaired performance averaged across all patients, and a proportion of 16.7% of the patients performing on an impaired level. However, separate scores for abstract thinking and cognitive flexibility were not reported, leaving it unclear whether the low performance was due to cognitive flexibility impairment alone. A cross-sectional study

found scores within normal limits on an abstract thinking test which does not involve cognitive flexibility, however large differences between patients were present[23]. One cross-sectional study reported impairments on a test measuring cognitive inhibition[21]. However, this result was based on two tests measuring information processing speed and cognitive inhibition. Therefore, it is unclear which process underlies the impaired performance.

Social cognition

Since none of the included studies examined social cognition with standardized neuropsychological tests, it remains unclear whether patients with CG are impaired in social cognition.

DISCUSSION

The current systematic review examined the incidence of cognitive impairment in patients with CG and reviewed the impairment in specific cognitive domains. Eleven studies were identified, including three case reports and one case-series. Of the eight studies investigating multiple patients, the quality was in seven studies moderate to low. Moreover, the number of studies per cognitive domain was low.

The review revealed that large differences exist amongst patients with CG. The averaged performance of the patients reported in each group study was often on a below average to low level, while a proportion of the patients performed on an impaired level. Twenty to 40 % of the patients performed on an impaired level on attention and memory, and, according to one study, on working memory. The range of vocabulary performances of the individual patients also exceeded the level of impairment, however specific percentages of the proportion of patients performing on an impaired level remained unknown. Evidence for impairments in other aspects of language functioning was mainly limited to case studies. The average level of performance did reach an impaired level for information processing speed, space perception, cognitive flexibility and cognitive inhibition, but the evidence was based on only a small number of studies. There is some indication that abstract thinking and visuoconstruction are relatively spared. Social cognition was not investigated at all. These results suggest that specific cognitive impairments indeed underlie the lower level of intellectual functioning. However, a specific cognitive profile cannot be determined due to individual differences between patients and limited number of merely small studies. A large number of studies investigating cognition in CG was excluded in this review since they only utilized developmental screening- or intelligence batteries. This represents the initial main focus on clinically assessing developmental delay and intelligence only in patients with CG.

To improve the knowledge concerning cognitive functioning in CG, well-designed and

well-reported studies covering multiple cognitive domains should be performed. The risk of selection bias needs to be lowered for example by describing the recruitment process in more detail and by refraining from the exclusion of patients with other long-term complications of CG (*i.e.* mental retardation and neurological or psychiatric disorders). Excluding these patients will lead to an underestimation of cognitive impairments that patients with CG may encounter. However, the inclusion of these patients will also add confounding factors, which will need to be taken into account in the statistical analysis. Moreover, the influence of late treatment on cognitive development[11,12] needs to be acknowledged by either solely including early-treated patients (*i.e.* < 8 weeks) or preferably distinguishing the results of early- and late treated patients. If available, the pathogenic mutations carried by the patients should be reported as well in order to establish the presence of patients with the homozygous p.Ser135Leu variation, which is associated with a better clinical outcome[13]. Possible confounders such as anxiety and depression[15] should be considered in the analysis of cognitive results, since they can influence cognitive performance[14] just as the level of intelligence[8]. Future neuropsychological assessments need to cover several cognitive domains with preferably multiple tests per domain. Important domains include information processing speed, attention, memory, language, visuospatial perception, executive functioning and social cognition. Information processing speed and executive functioning are especially important to be investigated in light of brain imaging findings in patients with CG. First, white matter abnormalities were found[29,30] which are associated with lower information processing speed in both healthy and patient populations[31]. Second, gray matter abnormalities were found within areas involved in executive functioning (*i.e.* medial prefrontal cortex and inferior frontal gyrus[32] and the orbital frontal lobes[33]). Finally, the report of cognitive results needs to involve the quantitative results of tests, including subtests. The mean raw score plus standard deviation and interquartile range, and normative score need to be reported as well as the proportion of patients performing on an impaired level to acknowledge the large individual differences between patients. Unfortunately, an important limitation of studying rare diseases is the absence of large sample sizes. In order to understand the differences in cognitive functioning between CG patients, large patient cohorts are necessary. Initiatives in which data of multiple patient cohorts are combined (e.g. the recently developed Galactosemia Patient Registry[34]) might help to facilitate studies investigating the relation between different disease parameters (e.g. variations in the *GALT* gene) and the severity of cognitive impairment in a larger group of patients. This will also make it possible to examine the association of cognition and other long-term complications of classical galactosemia such as movement disorders (e.g. tremor, ataxia, dystonia). Lower intellectual functioning has been found to be more frequent in patients with motor dysfunction[35], raising the question whether cognitive impairment in patients with CG is associated with other long-term complications of CG. Therefore, cognitive functioning should be part of this registry as well.

The recent international clinical guideline for the management of classical galactosemia[3] acknowledges that certain cognitive domains (*i.e.* executive functioning, information processing speed and visual spatial comprehension) need to be clinically assessed, additionally to the routine assessment of general mental abilities utilizing intelligence tests. This review supports this recommendation, but also highlights that neuropsychological assessment of CG patients should not be limited to these three cognitive domains. Preferably, all cognitive domains should be assessed by means of a neuropsychological assessment. In this way, patients and their caregivers will gain more insight in the patients' cognitive strengths and weaknesses. This will result in a better understanding by the patient, caregivers, teachers and coworkers and consequently a more suitable guidance plan can be made and access to appropriate interventions (e.g. compensatory strategy training[36]) can be provided. Ultimately, this could improve the health-related quality of life, which is affected by the lower level of cognitive functioning[5].

Methodological limitations

The review might have suffered from selection bias due to the inclusion of English, original articles only, and the exclusion of studies in which the diagnosis CG remained unclear or a specific sample of CG patients was drawn based on clinical outcome. This could have eliminated studies investigating cognitive functioning with standardized tests. However, only seven studies were excluded because of the latter two reasons, keeping the current sample of patients with CG representative of the entire CG population. Strengths of the current review include an extensive search strategy to incorporate all studies related to the long-term outcome of CG, and the inclusion of studies reporting quantitative data of standardized neuropsychological tests only.

CONCLUSION

This systematic review revealed that a large proportion of the patients (*i.e.* 20–40%) seems to perform on an impaired level on attention, memory and/or vocabulary. Evidence for impairments in information processing speed, language, visuospatial functioning, working memory, cognitive flexibility and cognitive inhibition was limited due to the small number of studies investigating these cognitive functions. Social cognition was not examined at all. Both clinicians and researchers encountering patients with CG need to be aware of possible cognitive impairments in different degrees of severity. However, they need to be conscious that only tentative conclusions regarding cognitive impairment can be drawn based on the current scientific evidence. All results need to be evaluated in larger, well-designed studies specifying the cognitive functioning and individual differences between CG patients in order to make a reliable judgement. This can be the basis for the development of intervention strategies.

ACKNOWLEDGEMENTS

The authors would like to thank Joost Daams, medical information specialist, for the development of the search strategies.

REFERENCES

1. Coelho AI, Rubio-Gozalbo ME, Vicente JB, Rivera I. Sweet and sour: an update on classic galactosemia. *Journal of inherited metabolic disease*. 2017;40(3):325–42.
2. Fridovich-Keil JL, Walter JH. Part 7: Carbohydrates, Chapter 72: Galactosemia. The Online Metabolic and Molecular Bases of Inherited Disease, OMMBID; Valle D.L., Antonarakis S, Ballabio A, Beaudet A.L., Mitchell G.A.(Eds.). McGraw Hill, New York.
3. Welling L, Bernstein LE, Berry GT, Burlina AB, Eyskens F, Gautschi M, et al. International clinical guideline for the management of classical galactosemia: diagnosis, treatment, and follow-up. *Journal of inherited metabolic disease*. 2016; 40(2):171–6.
4. Welling L, Waisbren SE, Antshel KM, Colhoun H-O, Gautschi M, Grünewald S, et al. Systematic review and meta-analysis of intelligence quotient in early treated individuals with classical galactosemia. *JIMD reports*. 2017;37:115–23.
5. Bosch AM, Grootenhuis MA, Bakker HD, Heijmans HS, Wijburg FA, Last BF. Living with classical galactosemia: health-related quality of life consequences. *Pediatrics*. 2004;113(5):e423–e8.
6. Welling L, Meester-Delver A, Derks TG, Janssen MC, Hollak CE, de Vries M, et al. The need for additional care in patients with classical galactosaemia. *Disability and Rehabilitation*. 2018:1–6.
7. American Psychiatric Association. Diagnostic and statistical manual of mental disorders (DSM-5). Washington: American Psychiatric Pub; 2013.
8. Lezak M, Howieson D, Bigler E, Tranel D. Neuropsychological assessment. New York, NY: Oxford University Press; 2012.
9. Moher D, Liberati A, Tetzlaff J, Altman DG, Prisma Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS medicine*. 2009;6(7):e1000097.
10. Van Eck N, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*. 2009;84(2):523–38.
11. Schweitzer S, Shin Y, Jakobs C, Brodehl J. Long-term outcome in 134 patients with galactosaemia. *European journal of pediatrics*. 1993;152(1):36–43.
12. Waggoner D, Buist N, Donnell G. Long-term prognosis in galactosaemia: results of a survey of 350 cases. *Journal of inherited metabolic disease*. 1990;13(6):802–18.
13. Lai K, Langley S, Singh R, Dembure P, Hjelm L, Elsas L II. A prevalent mutation for galactosemia among black Americans. *Journal of pediatrics*. 1996; 128(1):89–95.
14. Castaneda AE, Tuulio-Henriksson A, Marttunen M, Suvisaari J, Lönnqvist J. A review on cognitive impairments in depressive and anxiety disorders with a focus on young adults. *Journal of affective disorders.* 2008;106(1–2):1–27.
15. Waisbren SE, Potter NL, Gordon CM, Green RC, Greenstein P, Gubbels CS, et al. The adult galactosemic phenotype. *Journal of inherited metabolic disease*. 2012;35(2): 279–86.
16. Strauss E, Sherman EM, Spreen O. A compendium of neuropsychological tests: administration, norms, and commentary. New York, NY: Oxford University Press; 2006.
17. Scottish Intercollegiate Guidelines Network. Critical appraisal notes and checklists http://www.sign.ac.uk/checklists-and-notes.html.

18. Joanna Briggs Institute. Critical appraisal checklists https://joannabriggs.org/ critical_appraisal_ tools. Accessed 03 Apr 2018.
19. Antshel KM, Epstein IO, Waisbren SE. Cognitive strengths and weaknesses in children and adolescents homozygous for the galactosemia Q188R mutation: a descriptive study. *Neuropsychology*. 2004;18(4):658–64.
20. Lewis FM, Coman DJ, Syrmis M, Kilcoyne S, Murdoch BE. Impaired language abilities and pre-linguistic communication skills in a child with a diagnosis of galactosaemia. *Early child development and care*. 2013;183(12):1747–57.
21. Doyle CM, Channon S, Orlowska D, Lee PJ. The neuropsychological profile of galactosaemia. *Journal of inherited metabolic disease*. 2010;33(5):603–9.
22. Hoffmann B, Wendel U, Schweitzer-Krantz S. Cross-sectional analysis of speech and cognitive performance in 32 patients with classic galactosemia. *Journal of inherited metabolic disease*. 2011;34(2):421–7.
23. Kaufman FR, McBride-Chang C, Manis FR, Wolff JA, Nelson MD. Cognitive functioning, neurologic status and brain imaging in classical galactosemia. *European journal of pediatrics*. 1995;154(2):S2–5.
24. Van Erven B, Jansma BM, Rubio-Gozalbo ME, Timmers I. Exploration of the brain in rest: resting-state functional MRI abnormalities in patients with classic galactosemia. *Scientific reports*. 2017;7(1):9095.
25. Lewis FM, Coman DJ, Syrmis M, Kilcoyne S, Murdoch BE. Differential phonological awareness skills in children with classic galactosemia: a descriptive study of four cases. *JIMD reports*. 2012;10:45–52.
26. Iakovou K, Kalogerakou M, Schulpis K. A patient with classical galactosemia is graduated with a university degree. *Journal of pediatric endocrinology and metabolism*. 2018; 31(10):1147–8.
27. Lewis FM, Coman DJ, Syrmis M, Kilcoyne S, Murdoch BE. Charting a seven year trajectory of language outcomes for a child with galactosemia. *Journal of developmental and behavioral pediatrics*. 2013;34(6):414–8.
28. Ng W, Xu Y, Wong L, Kaufman F, Buist N, Donnell G. Two adult galactosaemia females with normal ovarian function and identical GALT mutations (Q188R/R333G). *Journal of inherited metabolic disease*. 2003;26(1):75–9.
29. Nelson M Jr, Wolff J, Cross C, Donnell G, Kaufman F. Galactosemia: evaluation with MR imaging. *Radiology.* 1992;184(1):255–61.
30. Timmers I, Zhang H, Bastiani M, Jansma BM, Roebroeck A, Rubio-Gozalbo ME. White matter microstructure pathology in classic galactosemia revealed by neurite orientation dispersion and density imaging. *Journal of inherited metabolic disease*. 2015;38(2):295–304.
31. Turken U, Whitfield-Gabrieli S, Bammer R, Baldo JV, Dronkers NF, Gabrieli JD. Cognitive processing speed and the structure of white matter pathways: convergent evidence from normal variation and lesion studies. *Neuroimage*. 2008;42(2):1032–44.

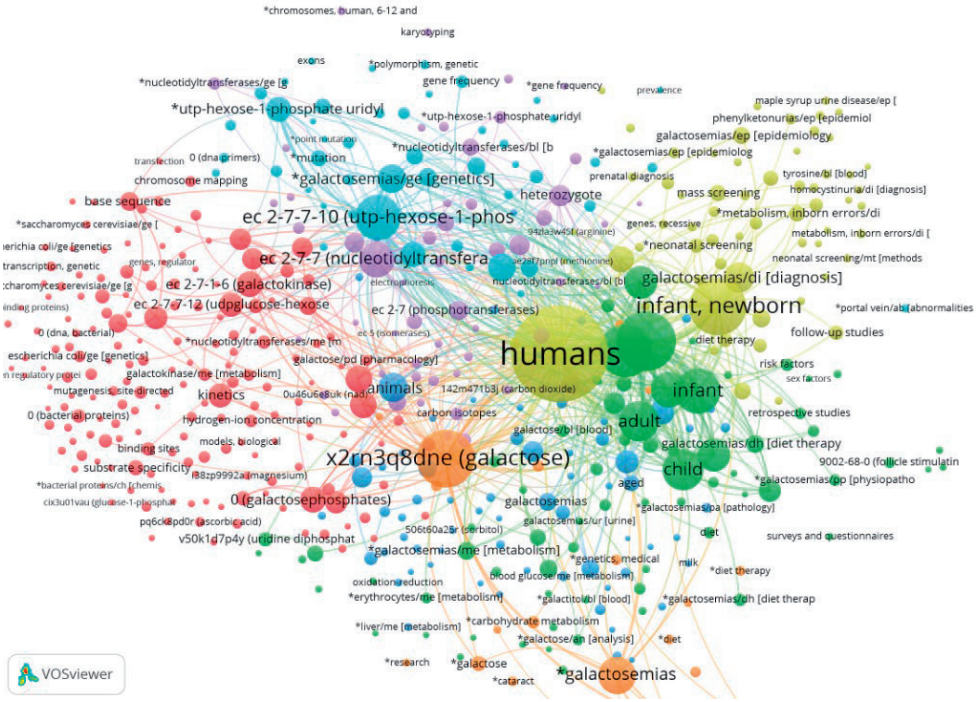
32. Timmers I, van der Korput LD, Jansma BM, Rubio-Gozalbo ME. Grey matter density decreases as well as increases in patients with classic galactosemia: a voxel-based morphometry study. *Brain research*. 1648;2016:339–44.
33. Dubroff J, Ficicioglu C, Segal S, Wintering N, Alavi A, Newberg A. FDG-PET findings in patients with galactosaemia. *Journal of inherited metabolic disease*. 2008; 31(4):533–9.
34. Rubio-Gozalbo ME, Bosch A, Burlina A, Berry G, Treacy E. The galactosemia network (GalNet). *Journal of inherited metabolic disease*. 2017;40(2):169–70.
35. Rubio-Agusti I, Carecchio M, Bhatia KP, Kojovic M, Parees I, Chandrashekar HS, et al. Movement disorders in adult patients with classical galactosemia. *Movement disorders*. 2013;28(6):804–10.
36. Fasotti L, Kovacs F, Eling PA, Brouwer WH. Time pressure management as a compensatory strategy training after closed head injury. *Neuropsychological rehabilitation*. 2000;10(1):47–65.

Supplementary Table 1. PRISMA checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4-5
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	-
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	5-6
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Add. Mat.
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	5-6
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	6-7
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	6
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	7
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	-

Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	7-8
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	-
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	-
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	8, Fig. 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	8-9, Tab. 1
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	9-10, Add. mat
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	Tab. 2
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	10-13
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	-
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	-
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	13-16
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	13-16
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	17
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	3

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097



Supplementary Figure 1. Clusters of related articles identified in VOSviewer.
Visualization of the identified clusters of related articles in VOSviewer which were used to manually omit irrelevant clusters of articles in the search strategies in EMBASE and MEDLINE (see **Supplementary Table 2)**.

Supplementary Table 2. Search strategies for MEDLINE, EMBASE and PsychINFO

1. Search strategy MEDLINE (Ovid MEDLINE(R) Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) <1946 to Present>)

#	Search terms
1	galactosemias/ or UTP-Hexose-1-Phosphate Uridylyltransferase/ or UDPglucose-Hexose-1-Phosphate Uridylyltransferase/
2	(galactos?emia? or GALT deficien* or (galactose adj2 phosphate) or "utp hexose 1 phosphate" or "UDPglucose-Hexose-1-Phosphate").ab,kf,ti.
3	("EC 2-7-7-10" or "EC 2-7-7-12" or "9016-11-9" or "9026-21-5").rn.
4	NTR2869.ab.
5	or/1-4
6	animals/ not humans/
7	(animal? or rat? or mouse or mice or rodent?).kf,ti.
8	6 or 7
9	5 not 8
10	(kinetics or base sequence or galactokinase or bacterial proteins or protein confirmation or crystallography or genetic transcription or chromosome mapping or udb glucose 4-epime).ab,kf,ti.
11	"ec 2.7.1.6".rn.
12	"ec 5.1.3.2".rn.
13	or/10-12
14	review.ab,kf,pt,ti.
15	13 not 14 [notting out VOS red]
16	(dna mutation analysis or polymerase chain reaction or exons or gene frequency).ab,kf,ti. [notting out VOS blue]
17	(cultured cells or clincal enzyme test* or nucleotidyltransfera* or phosphotransfera*).ab,kf,ti.
18	"ec 2.7.7".rn.
19	"ec 2.7".rn.
20	or/17-19 [notting out VOS purple]
21	9 not (15 or 16 or 20)
22	remove duplicates from 21

2. Search strategy EMBASE (Ovid Embase Classic+Embase 1947 to 2018 October 22)

#	Search terms
1	galactose 1 phosphate uridylyltransferase/ or hexose 1 phosphate uridylyltransferase/ or galactosemia/
2	(galactos?emia? or GALT deficien* or (galactose adj2 phosphate) or "utp hexose 1 phosphate" or "UDPglucose-Hexose-1-Phosphate").ab,kw,ti.
3	("EC 2.7.7.10" or "EC 2.7.7.12").ab,ez. [EC numbers]
4	("9016-11-9" or "9026-21-5").rn. [CAS numbers]
5	NTR2869.ab,cn.
6	or/1-5
7	(animal/ or animal experiment/ or animal model/ or nonhuman/) not human/
8	(animal? or rat? or mouse or mice or rodent?).kw,ti.
9	7 or 8
10	6 not 9
11	(kinetics or base sequence or galactokinase or bacterial proteins or protein confirmation or crystallography or genetic transcription or chromosome mapping or udb glucose 4-epime).ab,kw,ti.
12	"ec 2.7.1.6".ab,ez.
13	"ec 5.1.3.2".ab,ez.
14	or/11-13
15	review.ab,kw,pt,ti.
16	14 not 15 [notting out VOS red]
17	(dna mutation analysis or polymerase chain reaction or exons or gene frequency).ab,kw,ti. [notting out VOS blue]
18	(cultured cells or clincal enzyme test* or nucleotidyltransfera* or phosphotransfera*).ab,kw,ti.
19	"ec 2.7.7".ab,ez.
20	"ec 2.7".ab,ez.
21	or/18-20 [notting out VOS purple]
22	10 not (16 or 17 or 21)
23	remove duplicates from 22

3. Search strategy PsychINFO (Ovid PsycINFO <1806 to October Week 3 2018>)

#	Search terms
1	(galactos?emia? or GALT deficien* or (galactose adj2 phosphate) or "utp hexose 1 phosphate" or "UDPglucose-Hexose-1-Phosphate").ab,id,ti.

Supplementary Table 3. Risk of bias assessment

1. Risk of bias assessment for case-control studies based on the SIGN checklist for case-control studies

Author (year)	Research question	Comparability population of cases and controls	Exclusion criteria	Participation rate	Similarity participants and non-participants	Clear definition of cases	Clear definition of controls	Blinding	Outcome assessment	Confounding	Report of results
Antshel et al. (2004)[19]	Low	Low	Low	Low	?	Low	Low	?	Low	Low	Low
Lewis et al. (2013)[20]	Low	?	?	?	High	Low	Low	?	Low	High	Low

Notes. Low: low risk of bias, High: high risk of bias, ?: unclear.

2. *Risk of bias assessment for cross-sectional studies based on the JBI critical appraisal checklist for prevalence studies*

Author (year)	Sample frame	Sampling procedure	Sample size	Description subjects and setting	Validity of methods for identification of disorder and condition	Outcome assessment	Response rate
Doyle et al. (2010)[21]	Low	?	Low	Low	Low	Low	?
Hoffmann et al. (2011)[22]	Low	?	Low	Low	?	Low	High
Kaufman et al. (1995)[23]	Low	?	Low	Low	Low	High	?
Van Erven et al. (2017)[24]	Low	?	High	Low	Low	?	?
Waisbren et al. (2012)[15]	Low	?	Low	Low	Low	?	?

Notes. Low: low risk of bias, High: high risk of bias, ?: unclear.

3. Risk of bias assessment for case series based on the JBI critical appraisal checklist for case series

Author (year)	Inclusion criteria	Measurement of condition	Validity of methods for identification of disorder	Consecutive inclusion	Complete inclusion	Description of demographics	Description of clinical information	Description of outcomes	Description of site/ clinic
Lewis et al. (2012)[25]	Low	Low	Low	High	High	Low	Low	Low	Low

Notes. Low: low risk of bias, High: high risk of bias, ?: unclear.

4. *Risk of bias assessment for case reports based on the JBI critical appraisal checklist for case reports*

Author (year)	Description of patients' demographic characteristics	Description of patients' history	Description of current clinical condition	Description of diagnostic tests, assessment methods and results	Take-away lessons
Iakovou et al. (2018)[26]	Low	High	High	Low	Low
Lewis et al. (2013)[27]	Low	Low	High	Low	Low
Ng et al. (2003)[28]	Low	High	High	High	Low

Notes. Low: low risk of bias, High: high risk of bias, ?: unclear.

Chapter 4

Classical galactosemia: neuropsychological and psychosocial functioning beyond intellectual abilities

M.M. Welsink-Karssies, K.J. Oostrom, M.E. Hermans, C.E.M. Hollak, M.C.H. Janssen, J.G. Langendonk, E. Oussoren, M.E. Rubio-Gozalbo, M.C. de Vries, G.J. Geurtsen, A.M. Bosch

Orphanet Journal of Rare Diseases. 2020 Feb;15(1):42.

ABSTRACT

Background

Despite early diagnosis and treatment, Classical Galactosemia (CG) patients frequently develop long-term complications, such as cognitive impairment. Available literature primarily reports on general intellectual abilities and shows a substantially lower Full Scale Intelligence Quotient (FSIQ) in CG patients than in the general population. Both problems in social functioning as well as internalizing problems are often reported in CG patients. The combination of intelligence, cognitive functioning, behavior and social functioning has not been studied systematically in CG patients.

Methods

To determine if CG patients demonstrate a specific neuropsychological and psychosocial profile, we investigated intelligence, functioning on multiple cognitive domains, behavior and social functioning with a comprehensive neuropsychological test battery and questionnaires (self- and proxy-reported).

Results

The data of 48 patients, aged 4–47 years are reported. FSIQ ranged from 45 to 103 (mean 77 ± 14). A negative correlation between age and FSIQ was demonstrated (p=0.037) which resulted directly from the inclusion of four young 'milder' patients detected by newborn screening (NBS) with an expected better clinical outcome. Compared with normative data, patients had significantly lower but highly variable scores on all cognitive domains, especially on tests requiring mental speed. In the context of the FSIQ, 43% of the cognitive test results exceeded IQ based expectations. Overall, the patients' scores on social functioning were in the normal range but internalizing problems were frequently reported. In our cohort, an early initiation of dietary treatment due to NBS or family screening did not result in a more favorable neuropsychological outcome.

Conclusion

In this study, we demonstrated that as a cohort, CG patients have a below average intelligence and impaired cognitive functioning without a distinctive neuropsychological profile. The effect of age on neurocognitive functioning should be assessed in longitudinal studies. Social functioning was not impaired, but patients may be at risk for internalizing problems. Considering the large variability in cognitive, behavioral and social functioning and the finding that cognitive outcomes may exceed IQ based expectations, an individual evaluation and follow-up is warranted in all CG patients to ensure timely support if needed.

INTRODUCTION

Classical Galactosemia (CG, OMIM 230400) is an autosomal recessive inborn error of galactose metabolism, caused by a deficiency of the enzyme galactose-1-phosphate uridylyltransferase (GALT, EC 2.7.7.12). The ingestion of galactose from breastmilk or infant formula in the first weeks of life causes critical illness in affected neonates. A lifelong galactose-restricted diet is the only available treatment which is lifesaving in the newborn period but does not prevent long-term complications such as cognitive impairment, speech- and language deficits and movement disorders[1-3]. A published systematic review and meta-analysis demonstrated a substantially lower Full Scale Intelligence Quotient (FSIQ) in CG patients compared with the general population with large individual differences and a FSIQ ranging from fully normal to severely impaired[4]. The cognitive outcome of CG patients reported in the literature is mainly based on intelligence tests that lead to IQ. However, IQ is principally a dimension of individual differences in overall cognitive functioning called 'general intelligence'. It is the ultimate resultant of underlying more specific abilities. These abilities are referred to as cognitive functioning and encompass domains such as information processing speed, attention, memory, visuospatial functioning and executive functioning. Previous studies reporting on cognitive functioning in CG patients demonstrated below average to low functioning on several cognitive domains[5-9]. However, the outcomes on the cognitive domains differed between studies and results must be interpreted with care because studies mostly addressed only one cognitive domain, used one single test per cognitive domain and/or included small cohorts. A recently published systemic review demonstrated large differences between patients, but also suggested that specific cognitive impairments may cause the lower level of intellectual functioning observed in CG patients[10]. In order to investigate this properly, a comprehensive neuropsychological assessment addressing multiple cognitive domains with multiple tests per domain should be performed.

Besides the cognitive difficulties, problems in social functioning such as difficulties in making friends and maintaining a stable relationship, as well as internalizing behavior problems are often reported in CG patients and affect quality of life[2,9,11-13]. It has been suggested that CG patients exhibit autistic traits, however this has not been studied systematically and should be investigated in combination with social and behavioral functioning.

In order to provide patients with optimal support, more insight into the neurocognitive, social and behavioral functioning of CG patients is warranted. The aim of this study was to investigate the neuropsychological functioning of CG patients by assessing the combination of general intelligence, cognitive functioning on multiple domains, social functioning and behavior in a well-documented cohort of pediatric, adolescent and adult patients with CG. The effect of an early initiation of dietary treatment on neuropsychological functioning will be evaluated as well.

METHODS

Patients and recruitment

All pediatric and adult patients with CG, visiting our multidisciplinary galactosemia expertise outpatient clinic, were offered a standardized neuropsychological assessment as part of patient care according to the International guideline for CG patients[14]. Adult patients (≥18 years) or parents of patients (<18 years) were asked to complete the Social Responsiveness Scale (SRS). CG patients who are treated in other metabolic centers but participated in research in our outpatient clinic and recently received a neuropsychological assessment, data were retrieved after informed consent and incorporated if admissible.

Inclusion and exclusion criteria

All patients with erythrocyte GALT activity <15% of the reference mean and/or two known pathogenic variations in the *GALT* gene were eligible for participation in this study.

All patients with a second genetic diagnosis influencing clinical outcome were excluded. The results on cognitive functioning of patients with a FSIQ below 50 were excluded. The SRS of adult patients with an unknown FSIQ or a FSIQ below 70 were excluded.

Neuropsychological measures

The comprehensive neuropsychological assessment is summarized in **Supplementary Table 1** and includes standardized tests that cover:

- General intelligence
- Cognitive functioning on the domains: learning and memory, visuospatial functioning, executive functioning and mental speed.
- Questionnaires (self- and proxy reported) on executive functioning, behavior and social functioning

In children, Wechsler's Verbal IQ (VIQ) and in adults Wechsler's Verbal Comprehension Index (hereafter VIQ) indicates verbal functioning. In children, Wechsler's Performance IQ (PIQ) and in adults Wechsler's Perceptual Reasoning Index (hereafter PIQ) indicates non-verbal functioning.

Since it has been demonstrated that the subscales Withdrawn / depressed, Social problems and Thought problems (WST) of the Child Behavior Checklist (CBCL) 6–18y can be used to assess social functioning[15,16], these subscales and its sum were evaluated as well.

Data collection

The results of the neuropsychological assessment and data on educational attainment were collected and stored in an electronic clinical report form in Castor Electronic Data Capture, a good clinical practice compliant data management system[17].

Statistical analysis

SPSS version 25.0 (SPSS Inc. Chicago, Illinois, USA) was used to perform all statistical analyses. Descriptive analyses included means and standard deviations if data were normally distributed and median and ranges in case of a non-normal distribution. General intelligence was standardized to an IQ-score (mean 100, *SD* 15). All scores on cognitive functioning tests and questionnaires were standardized to T-scores (mean 50, *SD* 10), except for the HADS (Hospital Anxiety and Depression Scale) which is scored according to a Likert scale (0–3) resulting in a total score and the Developmental NEuroPSYchological Assessment (NEPSY) scores, which are expressed as percentile rank (pr) scores, ranging from well below the reference level (pr < 2) to above the reference level (pr > 75). Depending on the cognitive tests used, the standardized T-scores were corrected for age and/or gender and in adults most cognitive tests were corrected for educational attainment as well. The T-scores of patients were compared to normative data as reflection of the general population (T-score 50) with the use of the nonparametric sign test. A preliminary analysis showed a large variability in the FSIQ of patients. The effect of the FSIQ on cognitive functioning was evaluated by dividing patients into three FSIQ groups; group 1: FSIQ 50–69, group 2: FSIQ 70–85 and group 3: FSIQ > 85. Hereafter, individual cognitive test results were re-evaluated in the context of the FSIQ. More specifically, the FSIQ was converted into a T-score and was compared to the T-scores on the cognitive tests. In case a patient scored -1 *SD* (T-score ≤ 10) beneath their expected T-score, the cognitive test score was considered 'worse than expected' and a score above +1 *SD* (T-score ≥ 10) was defined as 'better than expected' based on the FSIQ. Differences between groups were tested with the Kruskal Wallis test or Mann-Whitney U test where appropriate. For the Hospital Anxiety and Depression Scale (HADS), the results of the patients were compared to the norm data of a reference group[18]. The Spearman's rank coefficient test was used to test for correlations and in case of a significant correlation, linear regression was performed. To evaluate the effect of possible confounders on our data, additional analyses were performed after the exclusion of patients with very limited data and patients with comorbidity potentially affecting cognitive functioning. To evaluate the effect of an early initiation of the diet on neurocognitive outcome, patients diagnosed before the introduction of newborn screening (pre-NBS) with a clinical presentation were compared to patients with an early diagnosis based on NBS or family screening (because of an older sibling with CG). P-values <0.05 were considered statistically significant. Multiple tests regarding a single hypothesis were corrected with the Bonferroni-Holm correction.

RESULTS

Of 67 CG patients visiting our multidisciplinary galactosemia expertise outpatient clinic and 6 CG patients treated in other metabolic centers, 54 patients received a neuropsychological assessment. Six patients were excluded because patients did not consent with the use of their clinical data for research purposes (*n*=3), had a second diagnosis influencing intellectual outcome (*n*=1), the administered tests were not part of our standardized assessment (*n*=1) and only a partly neuropsychological assessment was available due to visual impairment (*n*=1).

Demographics

In total, data of 48 patients are reported and demographics are presented in **Table 1**. The GALT erythrocyte activity was unknown in six patients with classical phenotypes. Our cohort includes four variant patients detected since the implementation of CG in the Dutch newborn screening (NBS) program in 2007 with residual erythrocyte GALT activity up 10% and possibly a better clinical outcome[19]. Two patients are homozygous for the p.Ser135Leu mutation with GALT deficiency in erythrocytes but residual GALT enzyme activity in other tissues which may improve clinical outcome[20]. The two homozygous p.Ser135Leu patients in our cohort were diagnosed late, at the age of 7 months and 10 years respectively. In the pre-NBS group (*n*=30) (diagnosis based on clinical symptoms), with the exception of the late diagnosed p.Ser135Leu patients, the diet was started at a median age of 10 days (6–39). In the early treated group (*n*=18) (diagnosis by NBS or family screening), the diet was started at a median age of 5.5 days (0–8).

Educational attainment

A total of 15 out of 48 patients (31%) attended or attend to date special schools for primary education compared to 4.5% in the general population[21] (**Table 1**). Of the patients aged 12 years and older who completed primary education, 9 out of 32 patients (28%) attend or attended special schools for secondary education compared to 3.0% in the general population[21]. In the Netherlands, one of the eligibility criteria for special education (smaller classes and tailored education) is a FSIQ below 80.

Of the patients who completed their education, 6 out of 15 patients (40%) completed education at a low educational level (of which five completed special education), 8 out of 15 patients (53%) completed education at the secondary vocational level and 1 out of 15 patients (7%) completed education at a high educational level, which is lower when compared with the general population (**Table 1**).

General intelligence

The FSIQ ranged from 45 to 103, with a mean of 77 (**Table 1**). The FSIQ did not significantly differ between males and females, nor between children and adults. Age

was significantly correlated with FSIQ (F(1, 46)=4.62, β0.42 (95%CI -0.82 – -0.03), *p*=0.037).
The results of the VIQ, PIQ and FSIQ of the Wechsler Scales of Intelligence are listed in **Table 1**. In 11 out of 48 patients, there was a significant difference of 15 or more points between the PIQ and VIQ. In seven patients (five adults and two children), this was in favor of the VIQ and the gap between VIQ and PIQ ranged from 15 to 24 IQ points. In four patients (all children), this was in favor of the PIQ and the gap between PIQ and VIQ ranged from 15 to 30 IQ points.

FSIQ and educational attainment
In 15 adult patients, the highest level of completed education was reported and in 5 out of 15 patients (33%) this was special education, while 13 out of 15 patients (85%) had a FSIQ below 80. Of the two patients with a FSIQ above 80, one patient completed secondary vocational education (FSIQ 81, 95%CI 77–87) and one patient completed higher professional education (FSIQ 88, 95%CI 83–93).

FSIQ and the initiation of treatment
There was no significant difference in VIQ, PIQ and FSIQ between patients in the pre-NBS group (*n*=30) and early treated patients (*n*=18). The exclusion of the late diagnosed p.Ser135Leu patients (*n*=2) and NBS-detected variant patients (*n*=4) did not change these results.

Cognitive functioning
The cognitive functioning test results are reported in **Table 2**. The reported T-scores of patients were compared to normative T-scores, based on a normative population sample.

Learning & Memory
On the AVLT, patients demonstrated lower scores on Immediate Recall, however the difference was not statistically different after correction for multiple testing. On the other two subtests, Delayed Recall and Delayed/Immediate Recall scores were comparable. On the Digit Span subtest, which requires auditory verbal memory and verbal working memory, patients had significantly lower scores.

Visuospatial functioning
The visuospatial functioning of patients was assessed with the GIT-2 spatial test and Block design. The significantly lowers scores for patients indicate a lower (visuo) spatial reasoning.

Table 1. Patient Demographics

Variable	All CG patients		
Gender, *n* (%)	Female: 28 (58) Male: 20 (42)		
Age (years)	Median 16 (4 – 47)		
GALT erythrocyte activity (%), *n* (%)			
- < 3.3	36 (75)		
- 3.3 – 8.7	6 (12.5)		
- Unknown	6 (12.5)		
Diagnosis based on (*n*):			
- Clinical symptoms (pre-NBS)	30		
- NBS	12		
- FS	6		
General Intelligence (*n*=48)			
- FSIQ	Mean 77 (95% CI 72 – 80), *SD* 14		
- VIQ	Mean 82 (95% CI 77 – 85), *SD* 15		
- PIQ	Mean 78 (95% CI 73 – 81), *SD* 15		
FSIQ, *n* (%)			
- FSIQ < 70	12 (25) (group 1)		
- FSIQ ≥ 70 – 85	20 (42) (group 2)		
- FSIQ > 85 – 100	15 (31) (group 3)		
- FSIQ > 100	1 (2) (group 3)		
Wechsler scales (*n*)	FSIQ	VIQ	PIQ
- WPPSI-IIINL (7)	Mean 84, *SD* 11	Mean 84, *SD* 17	Mean 94, *SD* 6
- WISC-IIINL (19)	Mean 79, *SD* 16	Mean 82, *SD* 16	Mean 76, *SD* 15
- WAIS-IVNL (22)	Mean 74, *SD* 13	Mean 81, *SD* 14	Mean 74, *SD* 13
Educational level, *n* (%)			
- Elementary school (*n*=48)			
- Normal education	25 (52)		
- Special education	15 (31)		
- Unknown	8 (17)		
- Secondary school (*n*=32)			
- Normal education	16 (50)		
- Special education	9 (28)		
- Unknown	7 (22)		
- Not applicable (age <12y)	16/48 (33)		
- Educational attainment patients*	*n*=17	Educational attainment Dutch population **	
1. Low educational level	35%	1.Low educational level	31.5%
2. Secondary educational level	47%	2. Secondary educational level	38.5%
3. High educational level	6%	3. High educational level	29%
Unknown	12%	Unknown	1.5%

Notes. CG: classical galactosemia, GALT: galactose-1-phosphate uridylyltransferase, NBS: newborn screening, FS: family screening, FSIQ: full scale IQ, VIQ: verbal IQ, PIQ: performal IQ, WPPSI: Wechsler Preschool and Primary Scale of Intelligence, WISC: Wechsler Intelligence Scale for Children, WAIS: Wechsler Adult Intelligence Scale. * Highest level of completed education, ** Data from the Dutch National Bureau of Statistics.

Executive functioning

- Inhibition

CG patients demonstrated significantly lower scores on the Stroop Inhibition, indicating a poor (response) inhibition. The comparable scores on the Stroop Interference indicate that patients do not have an increased sensitivity to interference.

- Cognitive Flexibility

Patients demonstrated lower scores on both the TMT B/A and Letterfluency. This indicates respectively that patients have an increased sensitivity to interference when it comes to cognitive flexibility and less flexibility in generating words. The comparable scores on all subtests of the WCST indicate that on a group level patients seem to be able to switch properly between strategies with a comparable amount of errors and perseverative responses.

Mental speed

The scores on TMT part A, which assesses visual and processing speed were comparable. The lower scores of patients on Stroop I & II, TMT part B indicate that patients needed more time to complete the tasks. On Symbol search and Substitution, which require processing speed, focusing attention and visual perception, patients had significantly lower scores.

Cognitive functioning: NEPSY results

Considering the results of the NEPSY are expressed as percentile rank scores, these results are reported separately and shown in **Table 3**. The NEPSY results of seven children are quite comparable to the results of the older patients reported above. On the NEPSY, approximately half of the patients demonstrated scores below the reference range on inhibition tasks and predominantly scores below the reference range on cognitive flexibility tasks. The mental speed was also impaired considering all patients needed more time to complete the tasks, which may indicate that inhibitory demands slow down cognitive processing speed. Although patients needed more time, all completed the NEPSY Naming and Switching tasks in the reference range. The scores on the subdomain attention indicate a poor selective and sustained attention.

Cognitive functioning and initiation of treatment

There was no significant difference in the cognitive functioning results between patients in the pre-NBS group and early treated patients. The exclusion of the homozygous p.Ser135Leu patients (n=2) did not change these results.

Table 2. Cognitive Functioning Results

Domain	*n*	Results patients	P-value*	*n*	FSIQ 50-69 (Group 1)	*n*	FSIQ 70-85 (Group 2)	*n*	FSIQ > 85 (Group 3)	P-value**
Learning & Memory										
- AVLT Immediate Recall	19	46.0 (9 - 61)	**0.029**	4	40.0 (9 – 46)	11	45.00 (35 – 61)	4	49.50 (47 – 51)	0.121
- AVLT Delayed Recall	19	46.0 (15 - 65)	0.545	4	49.5 (15 – 58)	11	46.00 (34 – 65)	4	50.00 (43 – 55)	0.956
- AVLT Delayed / Immediate	19	52.0 (38 - 65)	0.445	4	55.5 (43 – 62)	11	52.00 (38 – 65)	4	52.00 (39 – 62)	0.803
- Digit span	35	43.0 (20 – 63)	**<0.0005*****	8	30.00 (20 – 57)	15	43.00 (27 – 60)	12	48.50 (33 – 63)	**0.017**
Visuospatial functioning										
GIT-2 spatial test	19	36.0 (23 – 50)	**<0.0005*****	4	26.5 (23 – 40)	11	35.0 (28 – 40)	4	41.5 (40 – 50)	**0.019**
Block design	42	38.5 (20 – 53)	**<0.0005*****	9	30.0 (20 – 40)	17	37.0 (27 – 50)	16	37.0 (33 – 53)	**<0.0005*****
Executive functioning										
Inhibition										
- Stroop III (Inhibition)	25	45.0 (20 - 56)	**0.003*****	6	27.0 (20 – 49)	13	48.00 (22 – 56)	6	46.0 (35 – 53)	0.078
- Stroop III/II (Interference)	25	49.0 (30 - 66)	0.537	6	43.0 (30 – 60)	13	50.00 (31 – 66)	6	47.5 (40 – 63)	0.642
Cognitive flexibility										
- WCST Total number of errors	24	50.5 (27 - 67)	0.988	6	46.0 (27 – 50)	12	51.50 (37 – 67)	6	52.0 (39 – 64)	0.134
- WCST Perseverative responses	24	51.0 (30 - 81)	0.626	6	46.0 (33 – 52)	12	53.00 (30 – 81)	6	52.0 (35 – 73)	0.278
- WCST Percent Conceptual Level Responses	24	49.5 (27 - 64)	0.951	6	48.0 (27 – 51)	12	52.00 (39 – 64)	6	51.0 (37 – 62)	0.270
- TMT B/A	25	44.0 (27 - 57)	**0.002*****	6	45.0 (27 – 50)	13	43.00 (27 – 57)	6	48.5 (40 – 57)	0.510
- Letter fluency	19	37.0 (27 - 67)	**0.001*****	4	31.0 (28 -38)	11	39.00 (31 – 67)	4	34.0 (27 – 56)	0.143
Mental Speed										
- Stroop I (Color naming)	25	40.0 (25 - 61)	**0.001*****	6	35.0 (25 – 43)	13	43.00 (35 – 61)	6	46.5 (33 – 55)	0.077
- Stroop II (Word reading)	25	37.0 (20 - 56)	**<0.0005*****	6	30.0 (20 – 40)	13	43.00 (20 – 56)	6	39.0 (33 – 56)	0.063
- TMT A (Digit sequencing)	25	52.0 (20 - 67)	0.352	6	50.5 (20 – 59)	13	56.00 (43 – 67)	6	47.5 (33 – 67)	0.173
- TMT B (Digit-Letter-Switching)	25	45.0 (20 - 58)	**0.003*****	6	33.5 (20 – 47)	13	46.00 (20 – 58)	6	45.0 (33 – 56)	0.111
- Symbol search	41	43.0 (20 – 67)	**<0.0005*****	9	23.0 (20 – 50)	17	40.00 (27 – 67)	15	47.0 (40 – 60)	**0.001*****
- Substitution	42	40.0 (23 – 57)	**<0.0005*****	9	30.0 (23 – 47)	17	40.00 (30 – 57)	16	43.0 (33 – 53)	**0.006**

Notes. Data reported in T-scores, median (ranges). * Patient data vs. normative data (T-score 50), ** Comparison between FSIQ groups, *** Significant after Bonferroni-Holm correction. FSIQ: Full Scale IQ, AVLT: Auditory Verbal Learning Test, GIT-II: Groninger Intelligentie Test 2, Stroop: Stroop Color Word Test, WCST: Wisconsin Card Sorting Test, TMT: Trail Making Test.

Table 3. NEPSY-II Results

NEPSY-II	*n*	Low percentile rank scores* CG patients (*n*)	Normal percentile rank scores** CG patients (*n*)
Executive functioning			
Inhibition			
- Naming Total Errors	7	≤2 (1), 3-10 (1), 11-25 (1)	4
- Inhibition Total Errors	7	≤2 (1), 11-25 (2)	4
- Switching Total Errors	5	3-10 (1), 11-25 (2)	2
Cognitive flexibility			
- Response Set, Total Correct	5	≤2 (1), 3-10 (1), 11-25 (1)	2
Attention			
- Auditory Attention, Total Correct	7	≤2 (2), 3-10 (1), 11-25 (2)	2
Mental Speed			
- Naming Total Completion Time	7	-	7
- Inhibition Total Completion Time	7	3-10 (2), 11-25 (5)	-
- Switching Total Completion Time	5	-	5

Notes. Data reported in percentile rank scores. CG: classical galactosemia.
*low percentile rank scores: ≤2: well below the reference level, 3-10: below the reference level, 11-25: borderline / just below the reference level,
** normal percentile rank scores: 26-75: the reference level, >75: above the reference level.

Cognitive functioning in relation to general intelligence

The cognitive test results were compared between patients with a low FSIQ (50–69), an intermediate FSIQ (70–85) and a normal FSIQ (>85) (**Table 2**). On a group level, no significant differences were found on cognitive functioning between the FSIQ groups except for Symbol search (domain mental speed) and Block design (domain visuospatial functioning).

To evaluate if patients performed as expected based on their intellectual abilities, cognitive functioning was also individually evaluated in the context of the FSIQ. The results of the cognitive tests of adult and pediatric patients are shown separately (**Tables 4** and **5**). Of the adult patients, 4 out of 19 patients had one and 1 out of 19 patients had two worse than expected test scores. All patients had at least one better than expected test score. Of the pediatric patients, 4 out of 17 patients had a worse than expected test score on at least one and at most three tests. A majority of the patients (12 out of 17) had at least one better than expected test result.

In total, 43% of the cognitive test results were better than expected when evaluated in the context of the FSIQ.

BRIEF (behavior rating inventory of executive function) questionnaire

Six parents completed the BRIEF-P questionnaire (data not shown). On the Behavioral Regulation Index (BRI), Metacognition Index (MI) and the Total scale, one parent reported T-scores in the subclinical range. Four adolescents completed the BRIEF questionnaire. On the BRI, MI and Total scale, one patient reported T-scores above 50,

but well below the clinical range of 65.
Eighteen adults completed the BRIEF-A questionnaire. The median T-scores on the MI and Total Scale were higher in patients compared with normative data, but these differences were not significant after correction for multiple testing (p=0.007). On the BRI, 7 out of 18 patients (39%) scored in the subclinical range and 1 out of 18 patients (6%) reached the clinical range. On the MI, 11 out of 18 patients (61%) scored in the subclinical range and 3 out of 18 patients (17%) in the clinical range. High scores on the MI were reported on all subdomains (initiative, memory, planning and organizing, task evaluation and tidiness). There was no significant correlation between the FSIQ and the BRI, MI and Total Scale on all used versions of the BRIEF (BRIEF-P, BRIEF and BRIEF-A).

Behavioral functioning

CBCL (child behavior checklist) 6–18 years questionnaire

Parents reported scores in the subclinical and clinical range of the Internalizing problems scale only (**Table 6**). On the subdomains of the Internalizing problems scale, 'withdrawn/ depressed' and 'somatic complaints', parents reported significantly higher scores compared with normative data. No significant correlation between any of the CBCL outcome scales and FSIQ was found.

YSR (youth self-report) questionnaire

A total of three adolescents completed the YSR questionnaire and did not report any problems on the Internalizing, Externalizing and Total problems scales.

HADS (hospital anxiety and depression) questionnaire

Eighteen adults, four males (22%) and 14 females (78%) with a median age of 25.5 years (18–47) completed the HADS (**Table 7**). The results of patients were compared to the norm data of 947 controls with a median age of 37 years (18–47)[18]. A higher percentage of patients reported scores indicative for an anxiety disorder and depression when compared with the reference group, but the difference was not statistically significant. All patients reporting a score of 8 or higher have a FSIQ between 70 and 85.

Table 4. Individual Results, Adult Patients

Patient→	1	2	3	4	5	6[a]	7	8	9	10	11	12	13	14	15	16	17	18	19
VIQ	103	101	72	100	77	90	86	72	83	64	68	81	81	70	66	86	83	93	81
PIQ	104	77	72	85	60	72	75	62	75	58	60	75	87	70	68	75	77	77	81
Test results (corrected for FSIQ):																			
Executive functioning																			
Inhibition																			
Stroop III																			
Stroop III/II																			
Cognitive flexibility																			
- WCST Total Errors								X											
- WCST PR								X											
- WCST Percent CLR								X											
- TMT B/A																			
- Letter fluency								X											
Learning & Memory																			
- AVLT Immediate Recall								X											
- AVLT Delayed Recall								X											
- AVLT Delayed/ Immediate								X											
Digit span																			
Mental Speed																			
- Stroop I																			
- Stroop II																			
- TMT A																			
- TMT B																			
- Symbol search																			
- Substitution																			
Visuospatial functioning																			
GIT-2 spatial test								X											
Block design																			
Questionnaire results (uncorrected for FSIQ):																			
HADS*						- -		X				-				- -		- -	
BRIEF-A*	-		-		-	-		X	-			-			-	-	-	- -	-
SRS-A*		X	X		X	-				X	X	-		X			-	-	

Notes. [a]p.Ser135Leu homozygous patient. VIQ: Verbal IQ, PIQ: Performal IQ, FSIQ: Full Scale IQ, Stroop: Stroop Color Word Test, WCST: Wisconsin Card Sorting Test, PR: Perseverative Responses, CLR: Conceptual Level Responses, TMT: Trail Making Test, AVLT: Auditory Verbal Learning Test, GIT-II: Groninger Intelligentie Test 2, HADS: Hospital Anxiety and Depression Scale, BRIEF: Behavior Rating Inventory of Executive Function, SRS: Social Responsiveness Scale.

X no test result, ■ : test result worse than expected, ▒ : test result better than expected, □ : test result as expected.

* X no result, - - T-score on total scale in clinical range, - T-score on total scale in subclinical range, □ : T-score on total scale within normal range.

Table 5. Individual Results, Pediatric Patients

Patient→	20	21[b]	22[b]	23	24[b]	25	26	27	28	29	30	31	32	33	34	35[a]	36
VIQ	106	80	96	55	85	98	81	60	90	102	106	92	74	72	81	70	95
PIQ	93	100	85	85	97	85	59	75	77	89	99	86	61	85	95	57	98
Test results (corrected for FSIQ):																	
Executive functioning																	
Inhibition																	
- NI: Naming Total Errors	X			X		X					X	X	X	X	X	X	X
- NI: Inhibition Total Errors	X			X		X					X	X	X	X	X	X	X
- NI: Switching Total Errors	X	X	X	X		X					X	X	X	X	X	X	X
- Stroop III	X	X	X	X	X	X	X	X	X	X	X						
- Stroop III/II	X	X	X	X	X	X	X	X	X	X	X						
Cognitive flexibility																	
- NEPSY RS, Total Correct	X	X	X	X		X						X	X	X	X	X	X
- WCST Total Errors	X	X	X	X	X	X	X	X	X	X	X						
- WCST PR	X	X	X	X	X	X	X	X	X	X	X						
- WCST Percent CLR	X	X	X	X	X	X	X	X	X	X	X						
- TMT B/A	X	X	X	X	X	X	X	X	X	X	X						
- Letter fluency	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Attention																	
- NEPSY AA, Total Correct	X			X		X					X	X	X	X	X	X	X
Learning & Memory																	
- AVLT Immediate Recall	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
- AVLT Delayed Recall	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
- AVLT Delayed/ Immediate	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
- Digit span	X	X		X	X												
Mental Speed																	
- NI: Naming TCT	X			X		X					X	X	X	X	X	X	X
- NI: Inhibition TCT	X			X		X					X	X	X	X	X	X	X
- NI: Switching TCT	X	X	X	X		X					X	X	X	X	X	X	X
Stroop I	X	X	X	X	X	X	X	X	X	X	X						
Stroop II	X	X	X	X	X	X	X	X	X	X	X						
- TMT A	X	X	X	X	X	X	X	X	X	X	X						
- TMT B	X	X	X	X	X	X	X	X	X	X	X						
- Symbol search																	
- Substitution																	

Visuospatial functioning																	
- GIT-2 spatial test	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
- Block design																	
Questionnaire results (uncorrected for FSIQ):																	
CBCL 6-18y*	X			- -	X	X		X					- -	-			
YSR*	X	X	X	X	X	X	X	X	X	X	X	X	X				X
BRIEF-P*	-			X	X	X	X	X				X	X	X	X	X	X
BRIEF*	X	X	X	X	X	X	X	X	X	X	X	X	X				-
SRS-2*		X		X				-		-							

Notes. [a] p.Ser135Leu homozygous patient, [b] Variant patient, VIQ: Verbal IQ, PIQ: Performal IQ, FSIQ: Full Scale IQ, NI: NEPSY Inhibition, Stroop: Stroop Color Word Test, RS: Response Set, AA: Auditory Attention, WCST: Wisconsin Card Sorting Test, PR: Perseverative responses, CLR: Conceptual Level Responses. TMT: Trail Making Test, AVLT: Auditory Verbal Learning Test, TCT: Total Completion Time, GIT-II: Groninger Intelligentie Test 2, CBCL 6–18y: Child Behavior Checklist 6-18 years, YSR: Youth Self Report, BRIEF: Behavior Rating Inventory of Executive Function, SRS: Social Responsiveness Scale.

X no test result, ■: test result worse than expected, ▒: test result better than expected, □: test result as expected.

* X no result, - - T-score on total scale in clinical range, - T-score on total scale in subclinical range, □: T-score on total scale within normal range.

Table 6. CBCL 6-18y Results: Internalizing and Externalizing problems and Social Functioning

CBCL 6-18y, syndrome scale	*n*	Results patients	Normative data	T ≥ 65 ≤ 68 (subclinical range)	T > 68 (clinical range)	P-value
CBCL	14					
- Internalizing problems		58.0 (33 - 72)	50	-	-	0.157
· Anxious/depressed		54.0 (50 - 74)		*n* =2	*n* =1	**0.012**
· Withdrawn / depressed		58.0 (50 - 68)		*n* =3	*n* =0	**0.005***
· Somatic complaints		54.5 (50 - 64)		*n* =0	*n* =0	**0.003***
- Externalizing problems		37.5 (33 - 62)		-	-	**0.018**
· Rule-Breaking behavior		50.0 (50 - 60)		*n* =0	*n* =0	0.066
· Aggressive behavior		50.0 (50 - 62)		*n* =0	*n* =0	**0.042**
- WST		171 (150 - 201)	150	*n* =2	*n* =0	**0.001***
· Withdrawn/ depressed		58.0 (50 - 68)	50	*n* =3	*n* =0	**0.005***
· Social problems		56.5 (50 - 69)		*n* =1	*n*=3	**0.002***
· Thought problems		51.0 (50 - 72)		*n* =1	*n*=1	**0.007***

Notes. Data reported in T-scores, median (range). CBCL 6–18y: Child Behavior Checklist 6-18 years, WST: Withdrawn / depressed, Social problems & Thought problems. * Significant after Bonferroni-Holm correction.

Table 7. HADS Results

HADS	*n*	Patients	Total score ≥ 8 < 11 (%)	Total score ≥ 11 (%)	*n*	Reference group	Total score ≥ 8 < 11 (%)	Total score ≥ 11 (%)	P- value
	18				947				
- Anxiety scale		4.50 (1 – 16)	*n*=1 (6%)	*n*=3 (17%)		4.00 (0 – 17)	*n*=109 (12%)	*n*=71 (7%)	0.157
- Depression scale		3.00 (0 – 13)	*n*=0 (0%)	*n*=1 (6%)		2.00 (1 – 19)	*n*=82 (9%)	*n*=38 (4%)	0.241

Notes. Data reported in median (range). HADS: Hospital Anxiety and Depression Scale.

Social functioning

SRS (social responsiveness scale)

In total, 38 patients completed the SRS (**Table 8**). Seven patients (18%, 95%CI 11–36) had a T-score ≥61 indicating a mild to moderate impaired social responsiveness. One patient (2.6%, 95%CI 0.5–13.5) had a T-score of ≥76 indicating a severe disruption of social interaction in everyday life. Overall, the scores of patients were comparable to the normative data and individuals with elevated scores were within the expected frequency (based on a normal T-distribution the expected frequency of a score ≥61 is 16% and of a score ≥76 is 0.6%). The differences in T-scores between children and adults and males and females were not statistically significant. There was no significant correlation between the FSIQ and SRS-2 and FSIQ and SRS-A.

CBCL 6–18 years, questionnaire

The T-scores on the subdomains 'withdrawn/ depressed', 'social problems' and 'thought problems', as well as its sum (WST) were significantly higher in patients when compared with normative data (**Table 6**). A minority of the parents reported scores in the subclinical

and clinical range (on at least one and at most three domains).

Table 8. SRS results

SRS	*n*	Results patients	Normative data	T≥ 61 ≤ 75	T> 75	P-value
SRS-2 (parent)	23					
- Total		50.00 (40 – 85)	50	*n*=3	*n*=1	0.223
SRS-A (adult)	15					
- Total		53.00 (39 – 75)	50	*n*=4	*n*=0	0.116
SRS (all)	38					
- Total		52.50 (39 – 85)	50	*n*=7	*n*=1	0.073

Notes. Data reported in T-scores, median (range). SRS: Social Responsiveness Scale

Possible confounders

To evaluate the effect of possible confounders, additional analyses were performed. Most pediatric patients underwent limited cognitive testing due to their age. Therefore, analyses were repeated without the data of patients who underwent the Wechsler Scale of Intelligence and just one additional cognitive test (*i.e.* the NEPSY), which did not change the previous reported results.

The exclusion of patients with comorbidity (dyslexia *n*=2, ADHD *n*=3, possible autism spectrum disorder *n*=2, neonatal meningitis *n*=2, skull fracture *n*=1) and the exclusion of patients with scores indicative for an anxiety disorder or depression, did not change the previously reported results.

DISCUSSION

In this study, we aimed to investigate general intelligence, cognitive functioning on multiple domains, behavioral and social functioning in patients with CG. The results of this study demonstrate that as a group, CG patients have a below-average intelligence with a FSIQ of most patients between 70 and 85. Compared to normative data, patients demonstrated an overall impaired cognitive functioning, especially on tests within the domains mental speed, executive- and visuospatial functioning.

In our cohort, mental speed and visuospatial functioning were the most frequently affected cognitive domains. In previous studies, information processing speed (mental speed) has also been reported to be impaired[7,8], but the finding that visuospatial functioning was affected contradicts with previous research[7-9]. Until now, the subdomains of the domain executive functioning were not investigated separately. In our cohort, overall scores indicated poor (response) inhibition and less (cognitive) flexibility in generating words, but patients were able to switch properly between tasks which requires cognitive flexibility as well. Considering patients in our cohort demonstrated less flexibility in generating words and both receptive and expressive language impairments have been reported in multiple case-studies, language should be further investigated. On the domain learning and memory, patients had more difficulty learning words, but once learned the immediate and delayed recall were relatively spared in our cohort. Previous studies indeed found no impairments on delayed and immediate recall[7-9], even though a subset of patients did perform on an impaired level[10]. In order to correctly interpreted the cognitive functioning results, it is important to evaluate them in the context of intellectual abilities. Considering the included patients demonstrated an overall below average intelligence and a correlation between IQ and cognitive performance has been demonstrated[22], the cognitive outcomes were compared between patients with a low FSIQ (50–69), an intermediate FSIQ (70–85) and a normal FSIQ (>85). Interestingly, no significant differences in cognitive functioning between the three groups were found with the exception of two tests (block design and symbol search) that directly contribute to the FSIQ itself.

Considering the large differences in cognitive functioning between patients within the same FSIQ group, individual test results on cognitive functioning were re-evaluated in the context of intellectual abilities, which revealed that patients' cognitive outcomes may exceed IQ based expectations. This emphasizes the general idea that the FSIQ is an ultimate resultant that may not reflect underlying specific qualities and vulnerabilities of the individual patient, and that a more comprehensive neuropsychological evaluation will provide a better insight in one's strengths and weaknesses. According to the literature, especially executive functioning is crucial for academic performance and has a predictive value for academic achievement[23], which could also explain the higher educational attainment in our patients than their FSIQ would suggest.

Even though we investigated cognitive functioning on multiple domains with multiple tests and evaluated scores in context of the intellectual abilities, the large variability in cognitive functioning remained and a clear profile could not be distinguished. This variability complicates the interpretation of results and makes it impossible to draw an overall conclusion on cognitive functioning in CG. This finding underlines the need for an individual assessment in all patients. The large intra-individual variability in cognitive functioning has been demonstrated in healthy adults as well and the question remains whether found abnormalities indeed indicate the presence of brain dysfunction[24]. Therefore results should be evaluated in the context of daily functioning of the individual patient.

The VIQ and PIQ were evaluated as well. Overall, patients demonstrated a slightly higher VIQ than PIQ, which is in line with previous studies[8,25]. Considering the broad standard deviations and confidence intervals and small differences between VIQ and PIQ on a group level, it cannot be concluded that verbal skills (as numerically measured by VIQ) are better than non-verbal skills (PIQ). In those patients with a significantly higher VIQ however, this could potentially lead to an overestimation of the patients' abilities due to relatively good verbal skills. This is a relevant finding because in daily life this may put patients at risk for excessive demands in relation to their more limited cognitive abilities.

On the BRIEF questionnaires, only a minority of the already few parents reported executive functioning problems in the subclinical and clinical range which is in contrast to a vast majority of the adult patients. This may be explained by the fact that the children of these parents perform relatively well.

On the CBCL, all reported scores in the subclinical and clinical range were reported on the internalizing problems scale. Interestingly, on the YSR self-report, adolescents reported no problems even though the parents of two out of three patients reported scores in the clinical range on the internalizing problems scale. This is in line with a previous study which demonstrated that parents reported more problems than children, who might not experience or recognize problems in their functioning[9]. On the subscales of the CBCL indicative for social functioning (WST), CG parents reported statistically significantly higher problem scores than population norms. However, only a minority of parents reported scores in the subclinical or clinical range. Scores on the SRS were comparable to normative data and elevated scores indicating problems in social functioning were only reported by a minority of the patients. Importantly, our study did not find increased levels of social irresponsiveness or features suggestive of autism in CG. This is an important finding with regard to the observation that CG patients would manifest autistic traits. Future research is needed to understand poor social functioning in CG other than from deficits in the autism spectrum. A recent study found impaired visual information processing and facial emotion recognition in CG patients, which might contribute to the difficulties in social interactions observed in patients[26].

On the HADS, CG patients reported higher scores on the anxiety and depression scales than the reference group. The fact that this difference was not statistically significant might be due to the small numbers of patients included in this study. Since CG patients may be at risk for anxiety and depression, routine screening is warranted.
All but one patient reporting problems on the BRIEF and all patients with elevated scores on the HADS had a FSIQ between 70 and 85, suggesting that these patients may be at risk to develop problems in everyday life. The fact that these problems were not reported by patients with a FSIQ below 70 could indicate that they do not experience problems, are not aware of problems or simply have more difficulty in expressing problems. Another explanation could be that these patients are protected and guided more in everyday life or that the questionnaires used are not suited for patients with lower intellectual abilities.
Besides patients with classic phenotypes, our cohort also included four NBS-detected variant patients and two patients with a homozygous p.Ser135Leu mutation. The four variant patients received limited testing due to their age. Even though these patients are still young, they demonstrated a FSIQ above 85 and scores on the cognitive tests in the normal range. Follow-up of these patients is needed before it can be concluded that these patients indeed have a better neurocognitive outcome than patients with classical phenotypes. Patients with a homozygous p.Ser135Leu genotype are expected to have residual GALT enzyme activity in different tissues, which may improve their outcome. However, our two patients had a FSIQ of 71 and 61 and below average to low scores on several cognitive tests. These two patients did not present with critical illness in the neonatal period and were diagnosed late which resulted in a prolonged exposure to galactose which might explain this finding.
In our cohort, an early initiation of the diet because of NBS or family screening did not result in a higher FSIQ nor higher scores on the cognitive tests. Since most of the early treated patients are young and therefore received limited neuropsychological testing, follow-up is warranted before definitive conclusions can be drawn.
At this time it is unclear whether neurocognitive functioning declines with age as was suggested by cross-sectional studies[8,27,28], but contradicted by longitudinal studies[5,25]. In our cross-sectional study, age was negatively correlated with FSIQ. The exclusion of the younger variant patients in our cohort with an expected better clinical outcome, resulted in a non-significant correlation. Moreover, neurocognitive decline should be assessed in longitudinal studies.

Limitations
Not all CG patients visiting our expertise outpatient clinic underwent a neuropsychological assessment. Even though the patient group who chose not to undergo a neuropsychological assessment contains both patients with normal cognitive functioning and patients with an impaired cognitive functioning reported in their

medical charts, this may cause selection bias.
The data presented in this paper should be interpreted with care because a small number of patients provides statistical challenges. In adults, most T-scores on cognitive functioning tests are corrected for educational attainment, which might favor the results of the patients since they perform on a lower education level compared with the general population. However, significantly lower scores were still shown in patients when compared with the general population. Since patients have a substantially lower FSIQ compared with the general population, the individual results on cognitive functioning tests were evaluated in the context of the FSIQ of patients. This is indeed somewhat superficial, as the FSIQ arises from subtests of cognitive abilities that correlate with one another, and with the neuropsychological tests. Therefore, this controlling for consistent variables gives rise to positive manifold and may overshadow relevant cognitive impairments. Also, an interesting observation is that general intelligence appears to account for a larger share of cognitive variance in individuals exhibiting lower intelligence (as measured by IQ or mental age) than in individuals exhibiting higher intelligence[29].
In this study we did not intend to investigate the constructs of cognitive functioning. In order to assess cognitive functioning on multiple domains, the categorization of the domains as proposed by other studies was used[10]. The division in domains is needed to properly investigate cognitive functioning, but it is important to be aware that cognitive functioning tests may overlap between domains.
Longitudinal studies are needed to investigate apparent age-related dynamic changes between the different scales of intelligence as measured by the Wechsler scales. Moreover, re-assessment with latest iterations of the Wechsler scales will provide practitioners and scientist with more conceptual and practical insight into the developmental processes and the complex concept of intelligence in CG.
Besides intelligence, cognitive functioning, behavior and social functioning, there are other factors such as adaptive skills, personal, family and environmental factors that influence functioning of individuals to a certain extent and lay outside the scope of this article.
Adult patients completed the questionnaires during the neuropsychological assessment, whereas most parents completed the questionnaires at home. This resulted in a limited number of returned questionnaires completed by parents. The self-reported questionnaires might be hard on patients with an intellectual deficiency. Especially the SRS was difficult for patients with a FSIQ below 70 and therefore the results of these patients were not reported.

Strengths

In this study, we included all patients irrespective of their expected neuropsychological outcome and excluded patients with a second (genetic) diagnosis, which could influence neuropsychological functioning. Therefore, this is not only a relatively large but also a

representative CG cohort.
The assessment of cognitive functioning on specific domains with the use of multiple tests per domain provides a more reliable insight into the neuropsychological functioning of CG patients, than when only one test per domain is administered. The results of pediatric and adult patients were combined where possible. Since most pediatric patients received limited testing due to their age, analyses were repeated after the exclusion of these patients. The exclusion of these patients did not change the results and did not provide a more distinctive neuropsychological profile.
Since comorbidity, such as ADHD, autism, neonatal meningitis and dyslexia may cause executive functioning impairment[30] and anxiety and depression might be related to cognitive impairment and executive dysfunctioning in particular[31], analyses were repeated after the exclusion of these patients, however this did not change the results.

CONCLUSION

The current study provides insights in general intelligence, functioning on multiple cognitive domains, behavior and social functioning of patients with CG. As a group, patients have a substantially lower IQ and impaired cognitive functioning when compared with the general population and may be at risk for internalizing problems. Importantly, individual differences are considerable and specific cognitive abilities may exceed expectations that are based on the IQ. Based on the findings of our study, an individual neuropsychological assessment including the evaluation of behavior and social functioning is advised in all CG patients. In order to provide patients with timely and optimal support, the results of the neuropsychological assessment should be evaluated in a broader context, which includes adaptive functioning, the support system, the educational level and the capacity of patients and should include follow-up. This to ensure patients can reach their full potential without being subjected to excessive cognitive and emotional demands.

ACKNOWLEDGEMENTS

The authors would like to thank Prof. Dr. A.J.J.M. Vingerhoets, clinical psychologist, for sharing the HADS norm data collected by the research panel of CentERdata, institute for data collection and research. The authors would also like to thank Dr. J.H. van der Lee, epidemiologist, for her contribution with regard to the statistical analyses.

REFERENCES

1. Bosch AM. Classical galactosaemia revisited. *Journal of inherited metabolic disease*. 2006;29(4): 516–25.
2. Waisbren SE, Potter NL, Gordon CM, Green RC, Greenstein P, Gubbels CS, et al. The adult galactosemic phenotype. *Journal of inherited metabolic disease*. 2012;35(2): 279–86.
3. Coss KP, Doran PP, Owoeye C, Codd MB, Hamid N, Mayne PD, et al. Classical Galactosaemia in Ireland: incidence, complications and outcomes of treatment. *Journal of inherited metabolic disease*. 2013;36(1):21–7.
4. Welling L, Waisbren SE, Antshel KM, Colhoun HO, Gautschi M, Grunewald S, et al. Systematic review and meta-analysis of intelligence quotient in early treated individuals with classical Galactosemia. *JIMD reports*. 2017;37:115–23.
5. Manis FR, Cohn LB, McBride-Chang C, Wolff JA, Kaufman FR. A longitudinal study of cognitive functioning in patients with classical galactosaemia, including a cohort treated with oral uridine. *Journal of inherited metabolic disease*. 1997;20(4): 549–55.
6. Lewis FM, Coman DJ, Syrmis M, Kilcoyne S, Murdoch BE. Differential phonological awareness skills in children with classic galactosemia: a descriptive study of four cases. *JIMD reports*. 2013;10:45–52.
7. Kaufman FR, McBride-Chang C, Manis FR, Wolff JA, Nelson MD. Cognitive functioning, neurologic status and brain imaging in classical galactosemia. *European journal of pediatrics*. 1995;154(7 Suppl 2):S2–5.
8. Doyle CM, Channon S, Orlowska D, Lee PJ. The neuropsychological profile of galactosaemia. *Journal of inherited metabolic disease*. 2010;33(5):603–9.
9. Antshel KM, Epstein IO, Waisbren SE. Cognitive strengths and weaknesses in children and adolescents homozygous for the galactosemia Q188R mutation: a descriptive study. *Neuropsychology*. 2004;18(4):658–64.
10. Hermans ME, Welsink-Karssies MM, Bosch AM, Oostrom KJ, Geurtsen GJ. Cognitive functioning in patients with classical galactosemia: a systematic review. *Orphanet journal of rare diseases*. 2019;14(1):226.
11. Bosch AM, Grootenhuis MA, Bakker HD, Heijmans HS, Wijburg FA, Last BF. Living with classical galactosemia: health-related quality of life consequences. *Pediatrics*. 2004;113(5):e423–8.
12. Welling L, Meester-Delver A, Derks TG, Janssen MCH, Hollak CEM, de Vries M, et al. The need for additional care in patients with classical galactosaemia. *Disability and rehabilitation*. 2019;41(22):2663–8.
13. Hoffmann B, Dragano N, Schweitzer-Krantz S. Living situation, occupation and health-related quality of life in adult patients with classic galactosemia. *Journal of inherited metabolic disease*. 2012;35(6):1051–8.
14. Welling L, Bernstein LE, Berry GT, Burlina AB, Eyskens F, Gautschi M, et al. International clinical guideline for the management of classical galactosemia: diagnosis, treatment, and follow-up. *Journal of inherited metabolic disease*. 2017;40(2):171–6.

15. Quaedackers L, van Gilst MM, van Mierlo P, Lammers GJ, Dhondt K, Amesz P, et al. Impaired social functioning in children with narcolepsy. *Sleep*. 2019;42(2).
16. Biederman J, Petty CR, Fried R, Wozniak J, Micco JA, Henin A, et al. Child behavior checklist clinical scales discriminate referred youth with autism spectrum disorder: a preliminary study. *Journal of developmental and behavioral pediatrics*. 2010;31(6):485–90.
17. Castor EDC. (2019). Castor Electronic Data Capture. https://castoredc.com.
18. Vingerhoets AJJM. Norm data HADS; 2012.
19. Welling L, Boelen A, Derks TG, Schielen PC, de Vries M, Williams M, et al. Nine years of newborn screening for classical galactosemia in the Netherlands: effectiveness of screening methods, and identification of patients with previously unreported phenotypes. *Molecular genetics and metabolism*. 2017; 120(3):223–8.
20. Lai K, Langley SD, Singh RH, Dembure PP, Hjelm LN, Elsas LJ 2nd. A prevalent mutation for galactosemia among black Americans. *Journal of pediatrics*. 1996;128(1):89–95.
21. Data from the Dutch National Bureau of Statistics (www.cbs.nl).
22. Diaz-Asper CM, Schretlen DJ, Pearlson GD. How well does IQ predict neuropsychological test performance in normal adults? *Journal of the international neuropsychological society*. 2004;10(1):82–90.
23. Huizinga M, Baeyens D, Burack JA. Editorial: executive function and education. *Frontiers in psychology*. 2018;9:1357.
24. Binder LM, Iverson GL, Brooks BL. To err is human: "abnormal" neuropsychological scores and variability are common in healthy adults. *Archives of clinical neuropsychology*. 2009;24(1):31–46.
25. Schadewaldt P, Hoffmann B, Hammen HW, Kamp G, Schweitzer-Krantz S, Wendel U. Longitudinal assessment of intellectual achievement in patients with classical galactosemia. *Pediatrics*. 2010;125(2):e374–81.
26. Korner M, Kalin S, Zweifel-Zehnder A, Fankhauser N, Nuoffer JM, Gautschi M. Deficits of facial emotion recognition and visual information processing in adult patients with classical galactosemia. *Orphanet journal of rare diseases*. 2019;14(1):56.
27. Waggoner DD, Buist NR, Donnell GN. Long-term prognosis in galactosaemia: results of a survey of 350 cases. *Journal of inherited metabolic disease*. 1990; 13(6):802–18.
28. Komrower GM, Lee DH. Long-term follow-up of galactosaemia. *Archives of disease in childhood.* 1970;45(241):367–73.
29. Jensen AR. The g factor and the design of education. Sternberg RJW, Wendy M, editors editor; 1998. p. 111–31.
30. Anderson P. Assessment and development of executive function (EF) during childhood. *Child neuropsychology*. 2002;8(2):71–82.
31. Castaneda AE, Tuulio-Henriksson A, Marttunen M, Suvisaari J, Lonnqvist J. A review on cognitive impairments in depressive and anxiety disorders with a focus on young adults. *Journal of affective disorders*. 2008;106(1–2):1–27.
32. Wechsler D., Hendriksen, J., Hurks, P. (2009). WPPSI-III-NL: *Wechsler Preschool and Primary Scale of Intelligence – Third Edition – Nederlandstalige bewerking. Afname en Scoringshandleiding.* Amsterdam: Pearson Assessment and Information BV.

33. Kort, W., Compaan, L., Bleichrodt, N., Resing, W.C.M., Schittekatte, M., Bosmans, M., Vermeir, G. (2005). WISC-III-NL: *Wechsler Intelligence Scale for Children. David Wechsler. Derde Editie NL. Handleiding*. Amsterdam: NIP Dienstencentrum.
34. Wechsler D. (2012). WAIS-IV-NL: *Wechsler Adult Intelligence Scale – Fourth Edition – Nederlandstalige bewerking. Afname en scoringshandleiding.* Amsterdam: Pearson Assessment and Information BV.
35. Zijlstra HP, Kingma A, Swaab H, Brouwer WH. NEPSY-II-NL Nederlandstalige bewerking. Technische handleiding. Amsterdam: Pearson Assessment and Information B.V; 2010.
36. Hammes JGW. De Stroop Kleur-Woord Test. Pearson Assessment and Information BV: Handleiding; 1971.
37. Heaton RK, Staff PAR. Wisconsin card sorting test: computer Version-4. Research edition. Odessa: Psychological Assessment Resources; 2003.
38. Reitan RM. Trail making test: manual for administration and scoring. Tempe: Reitan Neuropsychology Laboratory; 1992.
39. Schmand B, Groenink SC, Van den Dungen M. Letterfluency:psychometrische eigenschappen en Nederlandse normen. Tijdschrift voor gerontologie en geriatrie. 2008;39(2):64–76.
40. Saan RJ, Deelman B. G. De 15-woordentest A en B (een voorlopige handleiding). Groningen: Afdeling Neuropsychologie, AZG; 1986.
41. Luteijn F, Barelds DPH. GIT2: Groninger Intelligentie Test 2. Amsterdam: Harcourt Test Publishers. 2004.
42. Huizinga M, Diana Smidts D. BRIEF Handleiding: Oudervragenlijst executieve functies voor 5- tot 18-jarigen. Amsterdam: Hogrefe; 2012.
43. Huizinga M, Smidts D. BRIEF Handleiding: Zelfrapportagelijst executieve functies voor 11- tot 18-jarigen. Amsterdam: Hogrefe; 2012.
44. Scholte E, Noens I. BRIEF-A Handleiding: Vragenlijst executieve functies voor volwassenen. Amsterdam: Hogrefe; 2011.
45. Verhulst F, Van der Ende J. Handleiding Gedragsvragenlijst voor kinderen (Child Behavior Checklist): CBCL 6-18 jaar oudervragenlijst. Achenbach System of Emperically Based Assessments Nederland; 1996.
46. Verhulst F, van der Ende J. Handleiding Gedragsvragenlijst voor kinderen (Child Behavior Checklist): YSR Youth Self-Report 11-18. Achenbach System of Emperically Based Assessments Nederland; 2000.
47. Van Hemert AM, Ormel J. Nederlandse Versie van de Hospital Anxiety and Depression Scale (HADS). Groningen: Rijksuniversiteit Groningen; 1996.
48. Constantino JN, Gruber C, Roeyers H, Thys M, Druart C, De Schryver M, et al. Handleiding SRS-2 Screeningslijst voor autismespectrumstoornissen. Amsterdam: Hogrefe Uitgevers BV; 2015.
49. Constantino JN, Gruber C, Noens I, De la Marche W, Scholte E. Handleiding SRS-A Screeningslijst voor autismespectrumstoornissen. Amsterdam: Hogrefe Uitgevers B.V; 2012.

Supplementary Table 1. The Neuropsychological Assessment

Age	2y;6m - 4y;11m	5y - 11y;11m	12y - 16y;11m	17y	≥ 18y
General intelligence (*n*)	- WPPSI-III^{NL} (2)	- WPPSI-III^{NL} (5) - WISC-III^{NL} (12)	WISC-III^{NL} (7)	WAIS-IV^{NL} (1)	WAIS-IV^{NL} (21)
Cognitive functioning (*tests*)		NEPSY-II	Stroop	Stroop	Stroop
			WCST	WCST	WCST
			TMT	TMT	TMT
				Letter Fluency	Letter Fluency
				AVLT	AVLT
			Digit span	Digit span	Digit span
		Symbol search	Symbol comparing	Symbol search	Symbol search
		Substitution	Substitution	Symbol substitution coding	Symbol substitution coding
				GIT-II spatial test	GIT-II spatial test
			Block design	Block design	Block design
	Block design	Block design			
Questionnaires					
Executive functioning		BRIEF-P (p)	BRIEF (c)	BRIEF (c)	BRIEF-A (a)
Behavior	CBCL 1,5-5 (p)	CBCL 1,5-5 (p)/	CBCL 6-18 (p)	CBCL 6-18 (p)	
		CBCL 6-18 (p)	YSR (c)	YSR (c)	
Anxiety & depression					HADS (a)
Social Functioning	SRS-2 (p) (≥ 4y)	SRS-2 (p)	SRS-2 (p)	SRS-2 (p)	SRS-A (a)

Notes. NL: Dutch version, y = years, m = months, p = reported by parents, c = self-reported by child, a = self-reported by adult

- General intelligence; the Wechsler Scales of Intelligence (WSI): the Wechsler Preschool and Primary Scale of Intelligence (WPPSI), completed by children ≥ 2y;6m until the 1st & 2nd grade of primary education[32], the Wechsler Intelligence Scale for Children (WISC), completed by children ≥ 3rd grade of primary education[33], the Wechsler Adult Intelligence Scale (WAIS)[34]. Variables of interest: Full Scale IQ (FSIQ), Verbal IQ (VIQ), Performal IQ (PIQ). Subtests used for cognitive functioning: Symbol search (symbol comparing), Substitution (symbol substitution coding), Digit span, Block design
- Cognitive functioning (tests); Developmental NEuroPSYchological Assessment (NEPSY)[35], Stroop Color Word Test (Stroop)[36], Wisconsin Card Sorting Test (WCST)[37], Trail Making Test (TMT)[38], Letter Fluency[39], Auditory Verbal Learning Test (AVLT)[40], GIT-II spatial test[41]
- Questionnaires (self- and proxy reported):
  - Executive functioning: BRIEF (Behavior Rating Inventory of Executive Function): BRIEF-P[42], BRIEF[43], BRIEF-A[44]
  - Behavior: CBCL (Child Behavior Checklist) 6-18y[45], YSR[46] and HADS (Hospital Anxiety and Depression Scale)[47]
  - Social Functioning: SRS (Social Responsiveness Scale): SRS-2[48], SRS-A[49]

PART 2

TOWARDS INDIVIDUAL PROGNOSTICATION

Chapter 5

The Galactose Index measured in fibroblasts of GALT deficient patients distinguishes variant patients detected by newborn screening from patients with classical phenotypes

M.M. Welsink-Karssies[*], M. van Weeghel[*], C.E.M. Hollak, H.L. Elfrink, M.C.H. Janssen, K. Lai, J.G. Langendonk, E. Oussoren, J.P.N. Ruiter, E.P. Treacy, M.C. de Vries, S. Ferdinandusse[#], A.M. Bosch[#]

[*] Authors contributed equally to this manuscript
[#] Authors contributed equally to this manuscript

Molecular Genetics and Metabolism. 2020 Mar;129(3):171-176.

ABSTRACT

Background

The high variability in clinical outcome of patients with Classical Galactosemia (CG) is poorly understood and underlines the importance of prognostic biomarkers, which are currently lacking. The aim of this study was to investigate if residual galactose metabolism capacity is associated with clinical and biochemical outcomes in CG patients with varying geno- and phenotypes.

Methods

Galactose Metabolite Profiling (GMP) was used to determine residual galactose metabolism in fibroblasts of CG patients. The association between the galactose index (GI) defined as the ratio of the measured metabolites [U^{13}C]Gal-1-P/ [$^{13}C_6$]UDP-galactose, and both intellectual and neurological outcome and galactose-1-phosphate (Gal-1-P) levels was investigated.

Results

GMP was performed in fibroblasts of 28 patients and 3 control subjects. The GI of the classical phenotype patients (n=22) was significantly higher than the GI of four variant patients detected by newborn screening (NBS) (p=0.002), two homozygous p.Ser135Leu patients (p=0.022) and three controls (p=0.006). In the classical phenotype patients, 13 out of 18 (72%) had a poor intellectual outcome (IQ<85) and 6 out of 12 (50%) had a movement disorder. All the NBS-detected variant patients (n=4) had a normal intellectual outcome (IQ≥85) and none of them has a movement disorder. In the classical phenotype patients, there was no significant difference in the GI between patients with a poor and normal clinical outcome. The NBS-detected variant patients had significantly lower GI levels and thus higher residual galactose metabolism than patients with classical phenotypes. There was a clear correlation between Gal-1-P levels in erythrocytes and the GI (p=0.001).

Conclusions

The GI was able to distinguish CG patients with varying geno- and phenotypes and correlated with Gal-1-P. The data of the NBS-detected variant patients demonstrated that a higher residual galactose metabolism may result in a more favorable clinical outcome. Further research is needed to enable individual prognostication and treatment in all CG patients.

INTRODUCTION

Classical Galactosemia (CG) is an autosomal recessive inborn error of galactose metabolism. Due to a deficiency of the galactose-1-phosphate uridylyltransferase enzyme (EC 2.7.7.12; GALT) patients are unable to metabolize galactose, which leads to the accumulation of galactose-1-phosphate (Gal-1-P) and galactitol. The only available treatment is a galactose-restricted diet. An early onset of the diet (after newborn screening (NBS) or family screening) improves neonatal outcome, but an early onset of and good compliance with the diet do not prevent long-term complications such as cognitive impairment, movement disorders and in females primary ovarian insufficiency[1-5]. The pathophysiology of CG and the broad clinical outcome spectrum ranging from fully normal to severely impaired are poorly understood. In patients with a GALT deficiency, Gal-1-P is persistently elevated even in dietary adherent patients, due to the endogenous production of galactose[6,7]. Both the accumulation of Gal-1-P and the reduced production of important substrates such as UDP sugars, are thought to contribute to the glycosylation defects demonstrated in CG patients and may thus contribute to the long-term complications[8-12]. At this time, there are no biomarkers that can predict the clinical outcome of CG patients. Prognostic biomarkers are urgently needed in all patients, but especially in patients with a classical phenotype in which clinical outcome varies highly and in patients detected since the implementation of NBS with previously unreported clinical and biochemical genotypes and phenotypes[13].

We hypothesize that differences in clinical outcome are caused by differences in residual galactose metabolism capacity. A slightly higher ability to metabolize galactose would cause lower Gal-1-P levels resulting in less abnormal galactosylation and possibly a more favorable clinical outcome. The method that is currently used in our cohort to measure erythrocyte GALT enzyme activity is not able to reliably detect differences in enzyme activity below 3.3% (<1.1 μmol/h.g Hb). To study the correlation between residual enzyme activity and clinical outcome, other methods should be investigated that are able to reliably detect (even slight) differences in residual galactose metabolism capacity. Metabolite profiling in fibroblasts has been demonstrated to correlate well with clinical severity in other metabolic disorders such as inborn errors of fatty acid oxidation[14,15]. In our pilot study we found that galactose metabolite profiling (GMP) in cultured fibroblasts was a sensitive method to determine residual galactose metabolism capacity expressed as the Galactose Index (GI). This method was able to differentiate between patients with a classical phenotype, patients with a variant phenotype and controls[16].

The aim of this study was to determine if residual galactose metabolism measured in fibroblasts can be used as predictor of clinical outcome by investigating the association between the GI and both biochemical and clinical outcomes in CG patients with varying genotypes and phenotypes.

METHODS

Patients and controls

Galactose metabolite profiling (GMP) was performed in fibroblasts of GALT deficient patients with two known pathogenic variations in the *GALT* gene and/or an erythrocyte GALT enzyme activity <15% of the reference mean and were collected for clinical or research purposes. The fibroblasts collected for clinical reasons were used after patients or their parents gave informed consent. The research fibroblasts were collected from competent, adult patients after informed consent was obtained. All patients consented to the use of their clinical data for research purposes. This study was approved by the local medical ethics committee. The fibroblasts of healthy controls were collected for research purposes after informed consent and approval of the local medical ethics committee.

Patient groups

This cohort comprises GALT deficient patients with varying geno- and phenotypes; patients with classical phenotypes (two pathogenic *GALT* mutations and absent or barely detectable erythrocyte GALT activity), NBS-detected variant patients (since 2007, with previously unreported geno- and phenotypes and erythrocyte GALT activity up to 10%, no clinical symptoms at diagnosis and undetectable Gal-1-P levels on dietary treatment)[13] and patients with the homozygous p.Ser135Leu genotype with residual GALT activity in other tissues than erythrocytes[17]. Patient demographics are listed in **Table 1**.

GMP measurements

The cell culture procedure and stable isotope ^{13}C labeled GMP measurements in fibroblasts were carried out as previously described[16]. In brief, fibroblasts were starved for 16 h in Dulbecco's phosphate buffered saline, followed by addition of 1 mM of [U^{13}C]Galactose. After an incubation period of 4 h metabolism was quenched and cells were processed for measurement of [U^{13}C]Galactose, [U^{13}C]Gal-1-P and [$^{13}C_6$]UDP-galactose. The galactose index (GI) was defined as the ratio of [U^{13}C]Gal-1-P/ [$^{13}C_6$] UDP-galactose. The GMP measurements were performed in triplicate and each patient cell line was measured in two independent experiments. Every experiment included at least two cell lines from control subjects and at least one cell line of a patient with a classical phenotype. The result (the GI) is the mean of the two independent experiments.

Clinical outcome

The clinical outcome parameters were intellectual and neurological outcome. In patients who received an age specific intelligence test, the IQ was used as derivative of intelligence. A poor intellectual outcome was defined as an IQ<85 and a normal intellectual outcome as an IQ≥85. The presence or absence of movement disorders (MDs) was used as a

determinant for neurological outcome. Information was retrieved from the medical records of the included patients.

Biochemical outcome

The most recent Gal-1-P level documented in the medical charts of patients was used in this study. All included Gal-1-P levels were measured by gas chromatography mass spectrometry (GC–MS) in erythrocytes and were below 0.82 μmol/g Hb in diet adherent patients. Patients with self-reported dietary incompliance at the most recent Gal-1-P measurement were excluded from the Gal-1-P analysis.

Table 1. Patient demographics and clinical outcomes

Pt ID	Pt ID*	Group	*GALT*_1 / *GALT*_2	GALT activity, %	GI	IQ	MD
1	12	C	p.Gln188Arg / p.Gln188Arg	< 3.3	14.27	81	Yes
2	14	C	p.Gln188Arg / p.Gln188Arg	< 3.3	18.53	78	-
3a	13	C	p.Gln188Arg / p.Gln188Arg	< 3.3	16.65	77	Yes
4a	8	C	p.Gln188Arg / p.Gln188Arg	< 3.3	15.58	71	-
5	10	C	p.Gln188Arg / p.Gln188Arg	< 3.3	22.37	83	-
6	4	C	p.Gln188Arg / p.Gln188Arg	-	13.94	91	-
7	9	C	p.Gln188Arg / p.Gln188Arg	-	11.49	53	No
8	7	C	p.Gln188Arg / p.Gln188Arg	< 3.3	17.49	82	No
9	-	C	p.Gln188Arg / p.Gln188Arg	< 3.3	14.00	-	-
10	-	C	p.Gln188Arg / p.Gln188Arg	< 3.3	13.16	-	-
11	-	C	p.Gln188Arg / p.Gln188Arg	< 3.3	10.12	-	No
12	-	C	p.Gln188Arg / p.Gln188Arg	< 3.3	10.77	-	No
13	31	C	p.Gln188Arg / p.Lys127E	< 3.3	18.98	70	-
14	21	C	p.Gln188Arg / p.Ser135Trp	< 3.3	13.49	98	No
15	27	C	p.Gln188Arg / p.Lys285Asn	< 3.3	16.18	77	-
16	34	C	p.Ser135Trp / p.Arg51Gln	< 3.3	12.87	78	-
17b	17	C	p.Gln188Arg / p.Leu195Pro	< 3.3	9.04	52	Yes
18b	19	C	p.Gln188Arg / p.Leu195Pro	< 3.3	11.08	88	Yes
19	30	C	p.Gln188Arg / p.Lys127E	< 3.3	13.76	61	-
20	18	C	p.Gln188Arg / p.Leu195Pro	< 3.3	16.71	93	No
21c	26	C	p.Gln188Arg / p.Lys285Asn	< 3.3	13.73	86	Yes
22c	25	C	p.Gln188Arg / p.Lys285Asn	< 3.3	16.54	76	Yes
23	36	S	p.Ser135Leu / p.Ser135Leu	3.9	5.40	61	No
24	37	S	p.Ser135Leu / p.Ser135Leu	< 3.3	2.47	71	No
25	38	V	p.Gln188Arg / p.Met219Lys	7.2	4.52	96	No
26	39	V	p.Gln188Arg / c.1-96T>G	3.6	6.31	86	No
27	-	V	p.Val128Ile / p.Val128Ile	9.3	7.17	91	No
28	-	V	p.Arg201His / p.Arg201His	8.9	1.61	-	-

Notes: *Patient (Pt) ID whole body galactose oxidation study[18]. C: Classical phenotype, S: homozygous p.Ser135Leu, V: Variant patients, IQ: Intelligence Quotient, MD: Movement Disorder. a,b,c: sibs. - : missing data.

Statistical analyses

SPSS version 25 (SPSS Inc. Chicago, Illinois, USA) was used to perform all statistical analyses. Data were presented as median and ranges because of a non-normal distribution. The Mann-Whitney U test was used to determine if statistically significant differences in the GI were present between patients with a poor and normal clinical outcome. The Spearman's rank coefficient test was used to test for associations and regression analysis was used to test for correlations. P-values <0.05 were considered statistically significant.

RESULTS

Patients

Fibroblast samples of 28 patients were included in this study. Of the included patients, 22 were classified as patients with a classical phenotype (12 patients with the homozygous p.Gln188Arg genotype) and four as NBS variant phenotype. Two patients were homozygous for the p.Ser135Leu mutation. Residual GALT enzyme activity in erythrocytes was below the limit of quantitation of the enzyme assay (<3.3%; <1.1 μmol/h.g Hb) in 21 out of 26 patients (**Table 1**). In two patients, the erythrocyte GALT enzyme activity was not documented.

GMP measurements

Galactose metabolite profiling (GMP) was performed in fibroblasts of 28 GALT deficient patients and three control subjects and the GI was calculated (**Table 1, Table 2 and Figure 1**). The GI of the classical phenotype patients was significantly higher than of the NBS-detected variants (p=0.002), the homozygous p.Ser135Leu patients (p=0.022) and controls (p=0.006). The GI of the NBS-detected variants was also significantly higher than of the controls (p=0.034). The GI of the homozygous p.Ser135Leu patients was increased compared with the GI of controls but the difference was not significant, probably due to the limited numbers (p=0.083). The GI was significantly lower in pediatric patients when compared with adult patients (p=0.030). In the classical phenotype group, there was no significant difference in the GI between pediatric and adult patients. There was no significant difference in the GI between males and females. The results of the individual GMP measurements including the coefficient of variation (CV) are listed in **Supplementary Table 1**. One experiment contained several outliers and therefore the entire experiment (the GMP measurements of 7 classical phenotype patients, 2 variants and 2 controls in triplicate) was excluded.

The association between the GI and clinical outcome

Our cohort included four NBS-detected variant patients with previously unreported genotypes and higher residual GALT enzyme activity than the patients with classical

phenotypes, and two homozygous p.Ser135Leu patients who were diagnosed at a later age (7 months and 10 years respectively) with residual GALT enzyme activity in other tissues. Since both the genotype and the late initiation of dietary treatment may influence clinical outcome, analyses were carried out both with and without these patients.

Table 2. Galactose Index (GI)

	n	Galactose Index
All patients	28	13.61 (1.61 – 22.37)
- Classical phenotypes	22	13.97 (9.04 – 22.37)
- NBS-detected variant patients	4	5.42 (1.61 – 7.17)
- Homozygous p.Ser135Leu	2	3.94 (2.47 – 5.40)
Controls	3	1.03 (0.84 – 1.12)

Notes. Data reported in median (ranges), NBS: newborn screening.

Intellectual outcome

The intellectual outcome of 24 patients was known, with an IQ ranging from 52 to 98 (median 78). The intellectual outcome was poor in 15 out of 24 patients (IQ<85) and normal in 9 out of 24 patients (IQ≥85). Overall, the GI in patients with a poor intellectual outcome was higher when compared with patients with a normal intellectual outcome, but the difference was not significant (p=0.053) (**Table 3**). After the homozygous p.Ser135Leu patients were excluded, the GI became significantly higher in patients with a poor intellectual outcome (p=0.010). In the group of classical phenotype patients, there was no significant difference in the GI between patients with a poor and normal intellectual outcome. Overall, there was no correlation between the IQ (as continuous variable) and GI (**Figure 1A**).

Neurological outcome

In our cohort, a movement disorder (MD; tremor and/or dystonia) was reported in 6 out of 17 patients. Overall, the GI in patients with an MD was higher when compared with patients without an MD, but the difference was not significant (p=0.132) (**Table 3**). The exclusion of the homozygous p.Ser135Leu patients did not change this finding. In the group of classical phenotype patients, there was no significant difference in the GI between patients with and without an MD.

The association between the GI and whole body galactose oxidation

This cohort includes 22 patients who have participated in a previous study, which investigated whole body galactose oxidation with the use of the 1-^{13}C galactose breath test[18].

Linear regression indicated a negative correlation between the GI and whole body

galactose oxidation (F(1,20) 30.47, β-0.89 (95%CI -1.22 – -0.55), p<0.0005) (**Figure 1B**). In the classical phenotype patients, there was no significant correlation between the GI and whole body galactose oxidation.

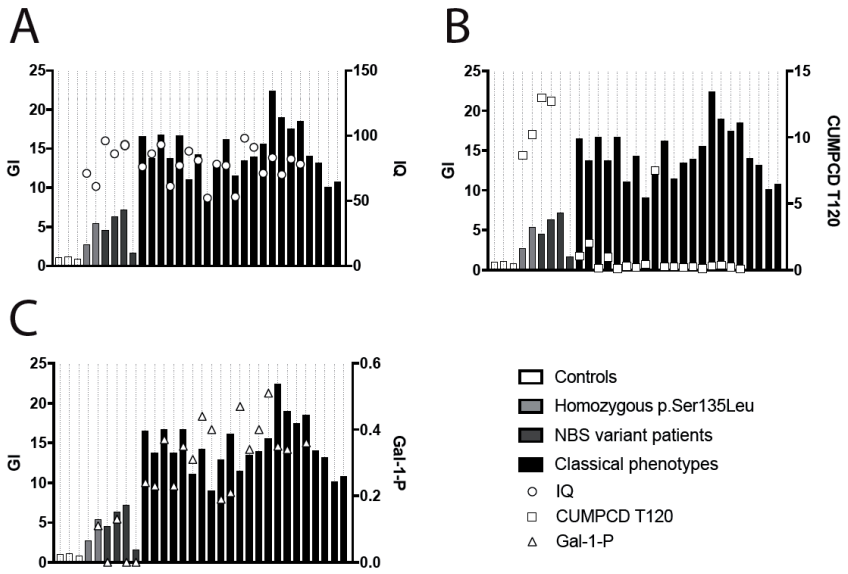


Figure 1. Individual data on the Galactose Index (GI) and IQ (1A), the GI and whole body galactose oxidation capacity after 120 minutes (CUMPCDT120) (1B) and the GI and galactose-1-phosphate (Gal-1-P) (1C) in classical galactosemia patients with varying phenotypes and controls

The association between the GI and biochemical outcome

For 21 out of 28 patients, the most recent Gal-1-P level measured by GC–MS was retrieved from the medical charts. Three NBS-detected variant patients demonstrated undetectable Gal-1-P levels (< 0.05 μmol/g Hb) under dietary treatment. Linear regression indicated a positive correlation between the GI and Gal-1-P (F(1,19) 13.89, β 2.19 (95%CI 0.96 – 3.43), p=0.001) (**Figure 1C**). In the group of classical phenotype patients, there was no significant correlation between Gal-1-P level and the GI.

Table 3. Galactose index and clinical outcome

	All patients		Classical phenotypes		NBS Variant patients		Homozygous p.Ser135Leu	
	n		*n*		*n*		*n*	
GI	28	13.49 (1.61 – 22.37)	22	13.97 (9.04 – 22.37)	4	5.42 (1.61 – 7.17)	2	3.94 (2.47 – 5.40)
- IQ ≥ 85:	9	11.08 (1.61 – 16.71)	5	13.73 (11.08 – 16.71)	4	5.42 (1.61 – 7.17)	-	
- IQ < 85:	15	15.58 (2.47 – 22.37)	13	16.18 (9.04 – 22.37)	-		2	3.94 (2.47 – 5.40)
- MDs, No:	11	10.12 (2.47 – 17.49)	6	12.49 (10.12 – 17.49)	4	5.42 (1.61 – 7.17)	2	3.94 (2.47 – 5.40)
- MDs, Yes:	6	14.00 (9.04 – 16.65)	6	14.00 (9.04 – 16.65)	-		-	

Notes. Data reported in median (ranges), GI: Galactose Index, IQ: Intelligence Quotient, MD: Movement Disorder.

DISCUSSION

In this study, we investigated the association between the GI measured in fibroblasts and the biochemical and clinical outcomes of GALT deficient patients. The previously developed GMP assay was used to determine residual galactose metabolism expressed as GI in a relatively large cohort of GALT deficient patients with varying geno- and phenotypes.

The results of this study demonstrate that the GI is able to differentiate between patients with a classical phenotype, NBS-detected variant patients with possibly better outcomes and controls, and results are in line with the conducted pilot study[16]. Furthermore, GMP analysis is able to distinguish homozygous p.Ser135Leu patients from classical phenotype patients and controls.

An important issue in CG is the lack of predictors of clinical outcome, especially in patients with a classical phenotype who have highly variable outcomes. In our cohort of classical phenotype patients, the differences in GI between patients with a poor and normal clinical outcome were not significant and thus the GI was not able to predict clinical outcome. As the GI demonstrated considerable interpatient and inter-assay variation within the classical phenotype patients with severely deficient GALT enzyme activity, the question remains whether future studies in a larger cohort would be able to demonstrate significant differences in GI between classical phenotype patients with a poor and normal clinical outcome.

The variant patients detected by NBS have residual GALT enzyme activities up to 10% in erythrocytes, significantly lower Gal-1-P levels than classical phenotype patients and currently none of them demonstrates long-term complications. The GI of these patients is significantly lower when compared with classical phenotype patients and their residual galactose metabolism capacity may prevent them from developing long-term complications, but as they are still young (below nine years of age) a careful long-term follow up is warranted.

The inclusion of the variant patients in our cohort resulted in a number of correlations.

Firstly, we found a significant correlation between age and GI, which was not found in our cohort of classical phenotype patients. Secondly, we found a clear correlation between GI and Gal-1-P levels, which is mainly due to the inclusion of the variant patients with higher residual galactose metabolism in fibroblasts and significantly lower Gal-1-P levels when compared with classical phenotype patients. Thirdly, the GI was negatively correlated with whole body galactose oxidation capacity, which can be attributed to the inclusion of variant patients who demonstrated lower GI levels and higher levels of whole body galactose oxidation than classical phenotype patients. Interestingly, within the group of classical phenotype patients with highly variable clinical outcomes, we found no correlation between the GI and both Gal-1-P levels and whole body galactose oxidation. In the classical phenotype patients, whole body galactose oxidation was consistently in the low range whereas the GI levels varied considerably. The differences between residual galactose metabolism in fibroblasts and whole body galactose oxidation might be attributed to organ specific GALT activity, which has been observed in an animal model[19]. The same study also demonstrated an age-dependent effect in different tissues. The results of the whole body galactose oxidation study indeed demonstrated significantly higher galactose oxidation capacity in younger patients, also within the group of classical phenotype patients. In the current study, GI levels were not affected by age within the group of classical phenotype patients. Therefore, both age and organ specific GALT activity may explain the differences between whole body galactose oxidation results and GI levels.

The finding that the variant patients demonstrated higher residual galactose metabolism in fibroblasts, lower Gal-1-P levels and possibly better clinical outcomes supports our hypothesis that differences in clinical outcomes are caused by differences in residual galactose metabolism capacity. The fact that we did not find this correlation in classical phenotype patients may be due to our small cohort or may suggest that the differences in clinical outcome in this subgroup may be attributed to other factors than residual galactose metabolism such as modifying genes.

There may be a threshold value for residual galactose metabolism in fibroblasts above which clinical outcome improves, but this could not be determined based on the results of this study.

A remarkable result are the low GI levels in our homozygous p.Ser135Leu patients with a poor intellectual outcome. Both patients also demonstrated whole body galactose oxidation in the control range and lower Gal-1-P levels than classical phenotype patients. Therefore, the question remains whether their poor intellectual outcome might be the result of the late initiation of dietary treatment as a comparable intellectual outcome has been reported in another late diagnosed homozygous p.Ser135Leu patient[12,20].

All variant patients were identified and treated early and none of them demonstrates long-term complications at this moment. Possibly their GI level provides sufficient galactose metabolism to protect against long-term complications. Based on the results of this study, the whole body galactose oxidation capacity and the undetectable Gal-1-P

levels in NBS-detected variants, some variant patients might even benefit from a less strict diet. Before dietary changes are implemented, the clinical outcomes of the variant patients need further follow-up and future studies into dietary relaxation are warranted.

Limitations

After the exclusion of outliers, the GMP measurements demonstrated an acceptable intra-assay variation, but the inter-assay variation was considerable especially in the group of classical phenotype patients. Despite additional analyses, a clear explanation for the variation could not be determined. The fluctuation in GMP measurements between experiments may indicate that differences in culture and/or incubation conditions influence the GMP measurements.

As previously described, the current method for GMP measurements is limited as it is not able to differentiate [$^{13}C_6$]-UDP-galactose from [$^{13}C_6$]-UDP-glucose and [U^{13}C]-Gal-1-phosphate from [U^{13}C]-glucose-1-phosphate[16].

Considering only labeled substrates are measured which have to pass the GALT enzyme first, this is only a small limitation. The question remains whether the observed differences in GI within the group of classical phenotype patients might be attributed to alternative disposal pathways.

CONCLUSION

The GI measured by GMP in GALT deficient fibroblasts distinguished patients with classical phenotypes from NBS-detected variants and homozygous p.Ser135Leu patients. In the classical phenotype patients, the GI was not able to differentiate between patients with a poor and normal clinical outcome. The data of the NBS-detected variants support our hypothesis that a higher residual galactose metabolism may result in a more favorable clinical outcome. Further research is needed to enable individual prognostication and treatment in all CG patients.

ACKNOWLEDGEMENTS

The authors would like to thank Evelien Tump and Esther Schneider for their assistance during the skin biopsies and Patricia Veltman and Petra Mooijer for cell culturing.

Funding

This study was supported by grants of The Galactosemia Foundation and Stichting Steun Emma.

REFERENCES

1. Komrower GM, Lee DH. Long-term follow-up of galactosaemia. *Archives of disease in childhood*. 1970;45(241):367-73.
2. Kaufman FR, Kogut MD, Donnell GN, Goebelsmann U, March C, Koch R. Hypergonadotropic hypogonadism in female patients with galactosemia. *New England journal of medicine*. 1981;304(17):994-8.
3. Bosch AM. Classical galactosaemia revisited. *Journal of inherited metabolic disease*. 2006;29(4):516-25.
4. Kuiper A, Grunewald S, Murphy E, Coenen MA, Eggink H, Zutt R, et al. Movement disorders and nonmotor neuropsychological symptoms in children and adults with classical galactosemia. *Journal of inherited metabolic disease*. 2019;42(3):451-8.
5. Rubio-Gozalbo ME, Haskovic M, Bosch AM, Burnyte B, Coelho AI, Cassiman D, et al. The natural history of classic galactosemia: lessons from the GalNet registry. *Orphanet journal of rare diseases*. 2019;14(1):86.
6. Berry GT, Nissim I, Lin Z, Mazur AT, Gibson JB, Segal S. Endogenous synthesis of galactose in normal men and patients with hereditary galactosaemia. *The Lancet*. 1995;346(8982):1073-4.
7. Huidekoper HH, Bosch AM, van der Crabben SN, Sauerwein HP, Ackermans MT, Wijburg FA. Short-term exogenous galactose supplementation does not influence rate of appearance of galactose in patients with classical galactosemia. *Molecular genetics and metabolism*. 2005;84(3):265-72.
8. Slepak TI, Tang M, Slepak VZ, Lai K. Involvement of endoplasmic reticulum stress in a novel Classic Galactosemia model. *Molecular genetics and metabolism*. 2007;92(1-2):78-87.
9. Fridovich-Keil JL, Walter JH. Part 7: Carbohydrates, Chapter 72: Galactosemia. The Online Metabolic and Molecular Bases of Inherited Disease, OMMBID; Valle D.L., Antonarakis S, Ballabio A, Beaudet A.L., Mitchell G.A.(Eds.). McGraw Hill, New York.
10. Lai K, Elsas LJ, Wierenga KJ. Galactose toxicity in animals. *International union of biochemistry and molecular biology (IUBMB) Life*. 2009;61(11):1063-74.
11. Coman DJ, Murray DW, Byrne JC, Rudd PM, Bagaglia PM, Doran PD, et al. Galactosemia, a single gene disorder with epigenetic consequences. *Pediatric research*. 2010;67(3):286-92.
12. Coss KP, Hawkes CP, Adamczyk B, Stockmann H, Crushell E, Saldova R, et al. *N*-glycan abnormalities in children with galactosemia. *Journal of proteome research*. 2014;13(2):385-94.
13. Welling L, Boelen A, Derks TG, Schielen PC, de Vries M, Williams M, et al. Nine years of newborn screening for classical galactosemia in the Netherlands: Effectiveness of screening methods, and identification of patients with previously unreported phenotypes. *Molecular genetics and metabolism*. 2017;120(3):223-8.
14. Diekman EF, Ferdinandusse S, van der Pol L, Waterham HR, Ruiter JP, Ijlst L, et al. Fatty acid oxidation flux predicts the clinical severity of VLCAD deficiency. *Genetics in medicine*. 2015;17(12):989-94.
15. van Rijt WJ, Ferdinandusse S, Giannopoulos P, Ruiter JPN, de Boer L, Bosch AM, et al. Prediction of disease severity in multiple acyl-CoA dehydrogenase deficiency: A retrospective and laboratory cohort study. *Journal of inherited metabolic disease*. 2019.

16. van Weeghel M, Welling L, Treacy EP, Wanders RJA, Ferdinandusse S, Bosch AM. Profiling of intracellular metabolites produced from galactose and its potential for galactosemia research. *Orphanet journal of rare diseases*. 2018;13(1):146.
17. Lai K, Langley SD, Singh RH, Dembure PP, Hjelm LN, Elsas LJ, 2nd. A prevalent mutation for galactosemia among black Americans. *Journal of pediatrics*. 1996;128(1):89-95.
18. Welsink-Karssies MM, van Harskamp D, Ferdinandusse S, Hollak CEM, Huidekoper HH, Janssen MCH, et al. The 1-^{13}C galactose breath test in GALT deficient patients distinguishes NBS-detected variant patients but does not predict outcome in classical phenotypes. *Journal of inherited metabolic disease*. 2019.
19. Coelho AI, Bierau J, Lindhout M, Achten J, Kramer BW, Rubio-Gozalbo ME. Classic Galactosemia: Study on the Late Prenatal Development of GALT Specific Activity in a Sheep Model. *The anatomical record: advances in integrative anatomy and evolutionary biology.* 2017;300(9):1570-5.
20. Crushell E, Chukwu J, Mayne P, Blatny J, Treacy EP. Negative screening tests in classical galactosaemia caused by S135L homozygosity. *Journal of inherited metabolic disease*. 2009;32(3):412-5.

Supplementary Table 1. GMP measurements

Pt ID	Exp. 1.1	Exp. 1.2	Exp. 1.3	Total GI Exp. 1	CV% Exp.1*	Exp. 2.1	Exp. 2.2	Exp. 2.3	Total GI Exp. 2	CV% Exp.2*	Total GI Exp. 1+2	Totale GI >2 Exp.
1	**5.88**	18.41	16.44	**13.58**/17.43	**50**/8	9.99	13.13	10.22	11.11	16	14.27	
2	15.26	20.21	16.75	17.40	12	16.99	20.90	21.08	19.66	12	18.53	
3a	15.80	7.50	26.64	16.65	**58**	7.85	16.29	31.28	18.47	**64**	16.65	
4a	20.69	22.63	21.49	21.60	5	14.05	11.47	10.71	12.08	15		15.58
	23.06	22.24	18.21	21.17	12	12.60	17.51	16.47	15.53	17		
	14.46	12.32	11.48	12.75	12	12.72	14.18	**24.42**	**17.11**/13.45	**37**/8		
	16.45	21.64	11.00	16.36	33	11.81	13.65	12.08	12.51	8		
5	22.37	23.90	23.51	23.26	3	20.78	21.24	22.41	21.48	4	22.37	
6	18.62	13.49	15.79	15.97	16	11.35	10.28	9.71	10.44	8		13.94
	22.48	22.99	23.89	23.12	3	15.28	18.39	12.58	15.42	19		
	11.37	11.70	11.27	11.45	2	14.30	11.33	10.51	12.05	17		
	8.35	7.21	**21.29**	12.28	**64**/10	10.46	9.92	7.07	9.15	20		
7	9.04	9.35	9.71	9.37	4	12.73	13.20	14.87	13.60	8	11.49	
8	19.01	18.96	20.19	19.39	4	18.98	10.89	16.89	15.59	27	17.49	
9	15.66	20.44	14.77	16.96	18	13.15	9.90	10.87	11.31	15	14.00	
10	16.53	14.44	**26.29**	**19.09**/15.49	**33**/10	10.93	10.55	11.02	10.83	2	13.16	
11	12.88	12.43	**25.56**	**16.95**/12.66	**44**/3	8.17	5.82	8.76	7.58	21	10.12	
12	12.77	12.47	11.59	12.28	5	8.77	9.15	9.84	9.25	6	10.77	
13	15.70	13.77	13.82	14.43	8	18.04	27.75	24.78	23.52	21	18.98	
14	13.18	11.52	**31.43**	**18.71**/12.35	**59**/10	17.73	13.41	12.76	14.63	18	13.49	
15	8.68	11.16	11.22	10.35	14	26.77	20.12	19.15	22.01	19	16.18	
16	14.11	15.99	13.61	14.57	9	13.03	11.14	9.33	11.17	17	12.87	
17b	11.98	9.62	8.76	10.12	16	7.36	7.91	8.58	7.95	8	9.04	
18b	14.89	11.30	14.62	13.60	15	7.26	8.94	9.47	8.56	13	11.08	
19	13.18	15.04	13.08	13.76	8	16.82	16.42	**39.11**	24.12	**54**/2	13.76	
20	17.53	15.78	16.81	16.71	5	22.76	13.55	9.53	15.28	**44**	16.71	
21c	12.83	13.78	14.56	13.73	6	**23.13**	28.05	27.57	26.25	10	13.73	
22c	14.61	17.34	17.68	16.54	10	**12.03**	29.55	22.35	21.31	**41**/20	16.54	
23	4.53	4.08	5.47	4.69	15	5.98	5.51	6.80	6.10	11	5.40	
24	**5.66**	3.22	2.50	**3.79**/2.86	**44**/18	2.14	1.96	2.14	2.08	5	2.47	
25	7.55	4.10	7.06	6.24	30	2.78	2.97	2.61	2.79	6	4.52	

26	5.37	7.06	6.51	6.31	14	**5.04**	12.09	11.91	9.68	**42**/1	6.31	
27	7.39	7.04	7.08	7.17	3	4.84	**8.56**	3.27	5.56	**49**/27	7.17	
28	1.58	2.34	1.60	1.84	24	1.40	1.38	1.34	1.37	2	1.61	
Co 01	0.69	0.55	0.59	0.61	11	0.93	1.02	0.76	0.90	15		1.03
	1.00	1.22	1.07	1.1	11	0.59	0.89	1.11	0.86	30		
	0.92	1.05	0.90	0.96	8	0.74	0.64	0.67	0.68	8		
	0.54	1.54	1.46	1.18	**47**/3	0.95	1.15	1.39	1.16	16		
	0.94	0.69	0.76	0.80	16	1.96	1.86	1.27	1.70	22		
Co 02	0.62	0.69	0.93	0.75	21	1.32	1.91	1.20	1.48	26	1.12	
Co 03	1.13	1.15	0.75	1.01	22	0.67	0.61	0.66	0.65	5		0.84
	1.15	1.03	1.01	1.06	7	0.76	0.88	0.67	0.77	14		
	1.44	0.58	0.64	**0.89**/ 0.61	**54**/7	0.73	0.53	0.66	0.68	17		
	1.01	0.68	0.89	0.86	19	1.36	1.14	1.32	1.27	9		

Notes. GMP: galactose metabolite profiling, a,b,c: sibs. Co: control subjects, Exp.: experiment, GI: Galactose Index, CV: coefficient of variation. * CV% without (bold)/ with the exclusion of outliers, : excluded measurement, bold: outlier

Chapter 6

The 1-^{13}C galactose breath test in GALT deficient patients distinguishes NBS-detected variant patients but does not predict outcome in classical phenotypes

M.M. Welsink-Karssies, D. van Harskamp, S. Ferdinandusse, C.E.M Hollak, H.H. Huidekoper, M.C.H. Janssen, E.M. Kemper, J.G. Langendonk, M.E. Rubio-Gozalbo, M.C. de Vries, F.A. Wijburg, H. Schierbeek, A.M. Bosch

Journal of Inherited Metabolic Diseases. 2020 May;43(3):507-517.

ABSTRACT

Background

Classical galactosemia (CG) patients frequently develop long-term complications despite early dietary treatment. The highly variable clinical outcome is poorly understood and a lack of prognostic biomarkers hampers individual prognostication and treatment. The aim of this study was to investigate the association between residual galactose oxidation capacity and clinical and biochemical outcomes in CG patients with varying geno- and phenotypes.

Methods

The noninvasive 1-^{13}C galactose breath test was used to assess whole body galactose oxidation capacity. Participants received a 7 mg/kg oral dose of 1-^{13}C labeled galactose. The galactose oxidation capacity was determined by calculating the cumulative percentage dose of the administered galactose (CUMPCD) recovered as $^{13}CO_2$ in exhaled air.

Results

Forty-one CG patients (5–47 years) and four adult controls were included. The median galactose oxidation capacity after 120 minutes (CUMPCDT120) of 34 classical patients (0.29; 0.08–7.51) was significantly lower when compared with two homozygous p.Ser135Leu patients (9.44; 8.66–10.22), one heterozygous p. Ser135Leu patient (18.59), four NBS-detected variant patients (13.79; 12.73–14.87) and four controls (9.29; 8.94–10.02). There was a clear correlation between Gal-1-P levels and CUMPCDT120 (p<0.0005).

Conclusion

In the classical patients, the differences in CUMPCDT120 were small and did not distinguish between patients with poor and normal clinical outcomes. The galactose breath test distinguished classical patients from homo- and heterozygous p.Ser135Leu and NBS-detected variant patients, but was not able to predict clinical outcomes in classical patients. Future studies are warranted to enable individualized prognostication and treatment, especially in NBS variants with galactose oxidation capacities in the control range.

INTRODUCTION

Classical Galactosemia (CG, OMIM 230400) is an inborn error of galactose metabolism, caused by a deficiency of galactose-1-phosphate uridylyltransferase (GALT, EC 2.7.7.12). Severe illness in the newborn period is prevented by an early start of dietary treatment after newborn screening (NBS) or family screening but long-term complications such as cognitive impairment, movement disorders (MDs) and in females ovarian failure are frequently seen in both screened and symptomatically detected patients[1-7]. The severity of long-term complications varies widely, even between siblings with identical mutations. Prognostic biomarkers are currently lacking, but are urgently needed for prognostication, especially in NBS-detected patients with previously unknown genotypes and phenotypes[8], and may facilitate individualized treatment.

The pathogenesis of CG and the cause of long-term complications is poorly understood. Endogenous galactose synthesis results in the persistent elevation of galactose1-phosphate (Gal-1-P) which is presumed to be toxic and thought to competitively inhibit other pathways such as the production of uridine diphosphate glucose (UDP) sugars, which are essential for the synthesis of glycoproteins and glycolipids[2,9,10]. Both accumulation of toxic metabolites and ongoing glycosylation abnormalities have been suggested to contribute to the long-term outcome in CG patients[11-13].

In search for prognostic biomarkers, we hypothesize that differences in clinical outcome are caused by differences in residual GALT enzyme activity. A slightly higher residual enzyme activity may lead to a clinically relevant higher galactose oxidation capacity. This will cause lower Gal-1-P levels resulting in less abnormal galactosylation and possibly a more favorable clinical outcome. While measurement of erythrocyte GALT activity is the gold standard for diagnosis, the method used in our cohort is not able to reliably detect differences in enzyme activity below 3.3% (<1.1 μmol/h.g Hb) and is therefore not suitable to study the correlation between residual enzyme activity and clinical outcome.

Previous studies showed that the non-invasive ^{13}C galactose breath test reflects individual galactose oxidation capacity and was able to differentiate between galactosemia patients and controls,[14-16] and even between *GALT* gene variations, such as the Duarte-variant (with erythrocyte GALT activity >15%) and p.Ser135Leu genotypes (with residual GALT activity in tissues other than erythrocytes)[15]. Although an association between galactose oxidation capacity and verbal dyspraxia has been found in CG patients[17], the association between residual galactose oxidation capacity and other clinical outcome measures, such as intellectual outcome and MDs have not been studied yet. Therefore, we determined the individual galactose oxidation capacity in our cohort of CG patients with different genotypes and phenotypes with the use of the ^{13}C galactose breath test and investigated the association with clinical and biochemical outcomes.

METHODS

Participants

CG patients were invited to participate in this study by their treating physician. Participating patients were studied in the galactosemia expertise outpatient clinic of the Amsterdam UMC.

CG patients with two known pathogenic variations in the *GALT* gene and/or an erythrocyte GALT activity <15% of the reference mean were eligible for participation in this study. This cohort comprises patients with varying geno- and phenotypes such as patients with classical phenotypes (two pathogenic *GALT* mutations and absent or barely detectable erythrocyte GALT activity), NBS-detected variant patients (detected since 2007 with previously unknown geno- and phenotypes[8]) and patients with the homo- and heterozygous p.Ser135Leu mutation[18]. The control group consisted of parents of pediatric patients with a confirmed heterozygous mutation in the *GALT* gene. Patients with swallowing difficulties and/or patients unable to follow breath test instructions were excluded.

Clinical outcomes

In order to investigate the association between galactose oxidation capacity and clinical outcomes, patients were divided into subgroups based on their intellectual and neurological outcome. Intellectual outcome was defined as poor (IQ<85) or normal (IQ≥85) and neurological outcome was based on the presence or absence of MDs.

Biochemical outcomes

The most recent Gal-1-P level measured by gas chromatography mass spectrometry (GC-MS) was used in this study (usual range in diet adherent patients 0.05–0.82 µmol/g Hb). Patients with self-reported dietary incompliance at the most recent Gal-1-P measurement were excluded from the Gal-1-P analysis.

^{13}C breath test

The 1-$^{13}C_1$ labeled galactose was produced by Cambridge Isotope Laboratories, Inc.©, Massachusetts, United States. Analysis in our laboratory showed a singly labeled stable isotope with a purity of 99%, as expected. To ensure a steady baseline of ^{13}C abundance in breath CO_2 over the study period, participants were instructed to eliminate ^{13}C enriched products from their diet 2 days prior to the start of the ^{13}C breath test[19] and to keep a galactose-restricted diet as specified in the international guideline[20] for at least 24 hours prior to the start of the study. This prevented a decreased enrichment of supplemented 1-^{13}C label with unlabeled galactose from the diet. A higher precursor enrichment (1-^{13}C galactose) will lead to higher $^{13}CO_2$ enrichment in breath. Participants were fasted for a minimum of 2 hours before start of the test and were only allowed to drink water during

the test. During the test, participants maintained a resting state. All participants received a 7 mg/kg oral dose of 1-^{13}C labeled galactose dissolved in water. Two baseline breath samples were collected before galactose ingestion. Patients were instructed to take a deep breath, to hold their breath for 3 seconds and to blow out into a glass vacutainer tube through a straw. Exactly 60, 90, and 120 minutes after the galactose ingestion, two breath samples were collected. Breath samples were stored at room temperature until analysis.

Measurement of $^{13}CO_2$ in breath samples

In each breath sample, the enrichment of ^{13}C in expired CO_2 was measured by automated gas-isotope-ratio mass spectrometry on a GC-IRMS Delta XL plus system (Thermo Fisher, Bremen, Germany) using a PoraBOND Q column (25 m × 0.32 mm; Agilent Technologies, the Netherlands)[21,22]. All collected samples were measured in triplicate. Standards and control samples were measured with each sequence of analyses. Results were expressed as the δ‰ vs the international reference standard, PeeDee Belemnite. Subsequently, CO_2 production was calculated with the use of the Schofield equation, based on gender, age, height, and weight of participants and the Weir equation using a fasting respiratory quotient (RQ) of 0.80[23]. The CO_2 production and the isotopic enrichment values were used to calculate the percentage dose recovered per participant at each time point and with these values the cumulative percentage dose of ^{13}C recovered from the labeled galactose (CUMPCD) as $^{13}CO_2$ in exhaled air was calculated[24]. The analytic variation was calculated for all participants.

Statistical analysis

SPSS version 25 (SPSS Inc. Chicago, Illinois, United States) was used to perform all statistical analyses. Descriptive analyses included median and ranges because of a non-normal distribution. To determine if there were statistically significant differences in CUMPCDT120 between patients with poor and normal clinical outcomes the Mann-Whitney U test was used. The Spearman's rank coefficient test was used to test for correlations and in case of a significant correlation, linear regression was performed. P-values <0.05 were considered statistically significant.

RESULTS

A total of 43 patients and four controls were included in this study. Two patients were excluded after being unable to follow breath test instructions due to their young age.

Demographics

The data of 41 patients, 19 males, and 22 females with a median age of 18 years (5–47) and four controls, one male and three females with a median age of 44 years

(38–49) are reported (**Table 1**). Our cohort consisted of 34 patients with a classical phenotype, hereafter classical patients, four NBS-detected variant patients, two patients with the homozygous p.Ser135Leu genotype and one heterozygous p.Ser135Leu/ p.*380Argext*50 patient. The erythrocyte GALT enzyme activity was below the limit of quantitation of the enzyme assay (<3.3%; <1.1 μmol/h.g Hb) in 32 out of 38 patients and between 3.6% and 8.7% in 6 out of 38 patients.

CUMPCDT120

The CUMPCD levels of the participants are summarized in **Table 2** and illustrated in **Figure 1**. The CUMPCD at 120 minutes (CUMPCDT120) of classical patients was significantly lower than of the NBS variants and of the homozygous p.Ser135Leu patients (p<0.019). The CUMPCDT120 was significantly higher in pediatric patients when compared with adult patients (p=0.001), which was also demonstrated in the classical pediatric patients (p=0.019). The CUMPCDT120 did not differ between males and females.

Correlation with age

Since the CUMPCDT120 was significantly higher in pediatric patients than in adult patients, the correlation between age and CUMPCDT120 was investigated in the largest group of patients with the same genotype, the homozygous p.Gln188Arg patients. This group consists of 14 classical patients with a median age of 23 years (6–35). Linear regression indicated a negative correlation between age and CUMPCDT120: F(1,12)5.77, β-0.30(95% CI -0,057 – -0,003), p=0.033.

Analytic variation

Since the galactose oxidation capacity of classical patients was found to be within a narrow range, the analytic variability of the galactose oxidation test was assessed. The effect of the analytic variation varied between patients with a low and high CUMPCDT120, and was ±0.06 in the patient with the lowest CUMPCDT120 of 0.08 and ±0.16 in the patient with the highest CUMPCDT120 of 18.59.

The association between CUMPCDT120 and clinical outcomes

The CUMPCDT120 analyses were performed for pediatric and adult patients separately (**Table 3**). The two included p.Ser135Leu homozygous patients were diagnosed late (7 months and 10 years), which may affect their clinical outcome. Therefore, analyses were carried out both with and without these patients. The youngest patient in our cohort with a p.Ser135Leu/ p.*380Argext*50 genotype demonstrated the highest CUMPCDT120 of 18.59 and seems an outlier. To evaluate the influence of this result, analysis were performed with and without this patient.

Table 1. Overview of included patients, demographics and galactose oxidation results

Pt ID	Group	*GALT*_1 / *GALT*_2	GALT activity, %	Age at diagnosis (days)	CUMPCD T120	IQ	MD
1	Cl	p.Gln188Arg / p.Gln188Arg	< 3.3	6	0.29	64	No
2	Cl	p.Gln188Arg / p.Gln188Arg	< 3.3	6	1.92	89	No
3	Cl	p.Gln188Arg / p.Gln188Arg	< 3.3	11	0.48	64	No
4	Cl	p.Gln188Arg / p.Gln188Arg	-	8	0.25	91	-
5	Cl	p.Gln188Arg / p.Gln188Arg	-	10	0.22	71	-
6a	Cl	p.Gln188Arg / p.Gln188Arg	< 3.3	0 (FS)	0.17	61	Yes
7	Cl	p.Gln188Arg / p.Gln188Arg	< 3.3	10	0.23	82	No
8b	Cl	p.Gln188Arg / p.Gln188Arg	< 3.3	0 (FS)	0.12	71	-
9	Cl	p.Gln188Arg / p.Gln188Arg	-	9	0.24	53	No
10	Cl	p.Gln188Arg / p.Gln188Arg	< 3.3	8	0.32	83	-
11	Cl	p.Gln188Arg / p.Gln188Arg	< 3.3	14	0.29	45	Yes
12a	Cl	p.Gln188Arg / p.Gln188Arg	< 3.3	-	0.21	81	Yes
13b	Cl	p.Gln188Arg / p.Gln188Arg	< 3.3	0 (FS)	0.12	77	Yes
14	Cl	p.Gln188Arg / p.Gln188Arg	< 3.3	14	0.08	78	-
15	Cl	p.Gln188Arg / p.Leu195Pro	< 3.3	9	4.59	82	No
16	Cl	p.Gln188Arg / p.Leu195Pro	< 3.3	10	0.28	103	Yes
17c	Cl	p.Gln188Arg / p.Leu195Pro	< 3.3	0 (FS)	0.43	52	Yes
18	Cl	p.Gln188Arg / p.Leu195Pro	< 3.3	18	0.14	93	No
19c	Cl	p.Gln188Arg / p.Leu195Pro	< 3.3	7	0.30	88	Yes
20	Cl	p.Gln188Arg / p.Ser135Trp	3.9	8	0.13	65	Yes
21	Cl	p.Gln188Arg / p.Ser135Trp	< 3.3	29	0.19	98	No
22	Cl	p.Gln188Arg / p.Ser135Trp	< 3.3	46	0.35	57	-
23	Cl	p.Gln188Arg / p.Ser135Trp	< 3.3	-	0.59	>85	-
24	Cl	p.Gln188Arg / p.Lys285Asn	< 3.3	22	0.15	49	Yes
25d	Cl	p.Gln188Arg / p.Lys285Asn	< 3.3	24	1.04	76	Yes
26d	Cl	p.Gln188Arg / p.Lys285Asn	< 3.3	24	2.02	86	Yes
27	Cl	p.Gln188Arg / p.Lys285Asn	< 3.3	-	0.24	77	-
28e	Cl	p.Arg148Gln / p.Trp316*	< 3.3	0 (FS)	0.53	68	Yes
29e	Cl	p.Arg148Gln / p.Trp316*	< 3.3	14	0.67	88	Yes
30	Cl	p.Gln188Arg / p.Lys127E	< 3.3	0 (FS)	0.96	61	-

31	Cl	p.Gln188Arg / p.Lys127E	< 3.3	-	0.37	70	-
32f	Cl	p.Gln188Arg / c.377+7A>C	< 3.3	0 (FS)	0.29	95	-
33f	Cl	p.Gln188Arg / c.377+7A>C	< 3.3	14	0.50	87	No
34	Cl	p.Ser135Trp / p.Arg51Gln	< 3.3	10	7.51	78	-
35	S	p.Ser135Leu / p.*380Rext*50	< 3.3	7	18.59	95	No
36	S	p.Ser135Leu / p.Ser135Leu	3.9	3860	10.22	61	No
37	S	p.Ser135Leu / p.Ser135Leu	< 3.3	210	8.66	71	No
38	V	p.Gln188Arg / p.Met219Lys	7.2	7	12.97	96	No
39	V	p.Gln188Arg / c.1-96T>G	3.6	8	12.73	86	No
40g	V	p.Val128Ile / p.Val128Ile	8.7	9	14.87	86	No
41g	V	p.Val128Ile / p.Val128Ile	8.7	9	14.61	89	No
42	Co	c.1-96T>G / -	-		9.27	-	-
43	Co	p.Gln188Arg / -	-		10.02	-	-
44	Co	p.Gln188Arg / -	-		8.94	-	-
45	Co	p.Ser135Trp / -	-		9.32	-	-

Notes: Cl: Classical, S: homo- and heterozygous p.Ser135Leu, V: Variants, Co: Controls, FS: Family Screening, IQ: intelligence quotient, MD: Movement Disorder. a,b,c,d,e,f,g: sibs, a&g: twins. - : missing data. Patients are listed on genotype and increasing age, : pediatric patients.

Table 2. CUMPCD levels of patients and controls

	n	CUMPCD T60	CUMPCD T90	CUMPCD T120
All patients	41	0.074 (-0.04 – 5.00)	0.19 (-0.006 – 11.09)	0.35 (0.08 – 18.59)
Controls	4	2.49 (2.26 – 2.62)	5.45 (5.23 – 5.84)	9.29 (8.94 – 10.02)
Classical patients	34	0.06 (-0.04 – 1.54)	0.16 (-0.006 – 3.87)	0.29 (0.08 – 7.51)
- Pediatric patients	14	0.10 (-0.04 – 0.58)	0.24 (-0.006 – 1.79)	0.49 (0.13 – 4.59)
- Adult patients	20	0.05 (-0.03 – 1.54)	0.13 (0.006 – 3.87)	0.24 (0.08 – 7.51)
NBS-detected variant patients	4	3.08 (2.33 – 3.49)	7.44 (6.12 – 8.25)	13.79 (12.73 – 14.87)
p.Ser135Leu homozygous	2	2.59 (2.39 – 2.79)	5.60 (5.14 – 6.07)	9.44 (8.66 – 10.22)
p.Ser135Leu/ p.*380Argext*50	1	5.00	11.09	18.59

Notes. Data reported in median and ranges, CUMPCD: Cumulative Percentage Dose Recovered.

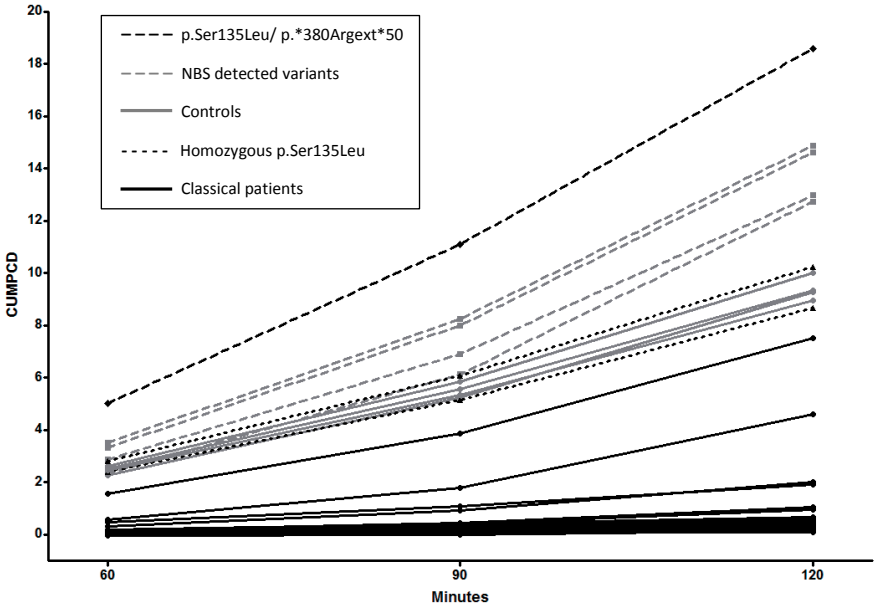


Figure 1. CUMPCD of patients and controls at 60, 90 and 120 minutes.

Intellectual outcome

In total, 8 out of 20 pediatric patients had a poor and 12 out of 20 had a normal intellectual outcome. Pediatric patients with a normal intellectual outcome demonstrated a significantly higher CUMPCDT120 than patients with a poor intellectual outcome (p=0.031). The difference remained significant after the exclusion of the homozygous p.Ser135Leu patient and the p.Ser135Leu/ p.*380Argext*50 patient. For the classical pediatric patients, there was no significant difference in CUMPCDT120 between patients with a poor and normal intellectual outcome.

In total, 16 out of 21 adult patients had a poor and 5 out of 21 had a normal intellectual outcome. There was no significant difference in CUMPCDT120 between adult patients with a poor and normal intellectual outcome. This result remained unchanged after the exclusion of the homozygous p.Ser135Leu patient.
In both pediatric and adult patients, there was no significant correlation between the IQ (as continuous measure) and CUMPCDT120.

Neurological outcome
An MD was found in 8 out of 19 pediatric patients and in 5 out of 10 adult patients. The CUMPCDT120 in pediatric patients without MDs was significantly higher when compared with pediatric patients with MDs (p=0.013). The difference remained significant after the exclusion of the homozygous p.Ser135Leu patient and the p.Ser135Leu/ p.*380Argext*50 patient. For the classical pediatric patients, there was no significant difference in CUMPCDT120 between patients with and without an MD. There was no significant difference in CUMPCDT120 between adult patients with and without an MD. These results remained unchanged after the exclusion of the homozygous p.Ser135Leu patient.

Siblings
Our cohort includes seven sibling pairs and their data are summarized in **Supplementary Table 1**.

The association between CUMPCDT120 and biochemical outcome
In 37 out of 41 patients, the most recent Gal-1-P level measured by GC-MS was available. Linear regression indicated a negative correlation between the CUMPCDT120 and Gal-1-P: F(1,36)33.43, β-0.21 (95%CI -0,029 – -0,014), p<0.0005. In the classical patients there was no significant correlation between CUMPCDT120 and Gal-1-P.

Table 3. CUMPCD T120 and clinical outcome in patient groups

	All patients (*n*=41)		Classical patients (*n*=34)		NBS variant patients (*n*=4)		Homozygous p.Ser135Leu (*n*=2)		p.Ser135Leu/ p.*380Argext*50 (*n*=1)	
Median age (years)		18.0 (5 – 47)		19.5 (6 – 47)		6.5 (5 – 8)		19.5 (16 – 23)		5.0
Median CUMPCD T120		0.35 (0.08 – 18.59)		0.29 (0.08 – 7.51)		13.79 (12.73 – 14.87)		9.44 (8.66 – 10.22)		18.59
Median CUMPCD T120 pediatric patients										
	n		*n*		*n*		*n*		*n*	
- IQ ≥ 85:	12	3.30 (0.28 – 18.59)	7	0.67 (0.28 – 4.59)	4	13.79 (12.73 – 14.87)	-		1	18.59
- IQ < 85:	8	0.45 (0.13 – 10.22)	7	0.43 (0.13 – 1.04)	-		1	10.22	-	
- MDs, No:	11	10.22 (0.29 – 18.59)	5	0.50 (0.29 – 4.59)	4	13.79 (12.73 – 14.87)	1	10.22	1	18.59
- MDs, Yes:	8	0.48 (0.13 – 2.02)	8	0.48 (0.13 – 2.02)	-		-		-	
Median CUMPCD T120 adult patients										
	n		*n*		*n*		*n*		*n*	
- IQ ≥ 85:	5	0.25 (0.14 – 0.59)	5	0.25 (0.14 – 0.59)	-		-		-	
- IQ < 85:	16	0.24 (0.08 – 8.66)	15	0.24 (0.08 – 7.51)			1	8.66		
- MDs, No:	5	0.23 (0.14 – 8.66)	4	0.21 (0.14 – 0.24)	-		1	8.66	-	
- MDs, Yes:	5	0.21 (0.12 – 0.30)	5	0.21 (0.12 – 0.30)			-			

Notes. CUMPCD: Cumulative Percentage Dose Recovered, NBS: newborn screening, IQ: Intelligence Quotient, MD: Movement Disorder.

DISCUSSION

The results of our study indicate that the ^{13}C breath test reflects individual galactose oxidation capacity by clearly distinguishing patients with classical phenotypes from NBS variants with higher erythrocyte GALT activity (up to 10%), and from homozygous p.Ser135Leu patients with deficient GALT activity in erythrocytes, but residual GALT activity in other tissues.

The results of most classical patients (CUMPCDT120 ≤ 2) and of the homozygous p.Ser135Leu patients (CUMPCDT120 in the same range as controls) were in line with previous research[15]. Since previous research demonstrated comparable galactose oxidation capacity between individuals heterozygous for p.Gln188Arg and healthy controls, we included heterozygous parents as controls[25].

On a group level, pediatric patients with a normal intellectual outcome and without MDs demonstrated a significantly higher galactose oxidation capacity than pediatric patients with a poor intellectual outcome and with MDs. This difference is caused by the inclusion of four NBS-detected variant patients in the pediatric cohort, whose galactose oxidation capacity even exceeded those of the controls. This might be due to their younger age, as we found a significant correlation between age and galactose oxidation capacity in patients with an identical genotype (homozygous p.Gln188Arg). Previous research also demonstrated that galactose oxidation capacity fluctuates in the first weeks of life[14] and in an animal model, age-dependent GALT activity was demonstrated in different tissues[26]. At this moment, it is unclear if younger patients truly have a higher galactose oxidation capacity or whether other factors that vary with age have influenced the breath test results, such as body composition, temporary label trapping[22] or differences in alternative disposal pathways, such as upregulation of the UDP-glucose pyrophosphorylase pathway.

The negative correlation found between the patients' CUMPCDT120 and Gal-1-P levels indicates that higher galactose oxidation capacity results in lower Gal-1-P levels. This is mainly the effect of the four NBS variant patients with higher galactose oxidation capacity and lower Gal-1-P levels than classical patients. Even though the NBS variant patients are young, currently none of them demonstrate long-term complications, which seems to support our hypothesis that differences in clinical outcome are caused by differences in residual galactose oxidation capacity. In the classical patients however, both CUMPCDT120 and Gal-1-P levels were within a narrow range and no significant differences in galactose oxidation capacity between patients with poor and normal clinical outcomes were demonstrated. This suggests that the differences in clinical outcomes in this subgroup may be attributed to other factors than differences in (residual) whole body galactose oxidation capacity.

In contrast to previous research in which all classical patients demonstrated CUMPCDT120 values < 2[15], two classical patients in our study had a CUMPCDT120 > 2. These patients have erythrocyte GALT activity below the limit of quantitation and demonstrated CG related illness at diagnosis. Since one patient is young with a classical

p.GlnQ188Arg/p.Leu195Pro genotype and a normal clinical outcome and the other is an older patient with a p.Ser135Thr/p.Arg51Gln genotype and a poor intellectual outcome, their results might be attributed to both age and genotype. This may also be the case for the youngest patient in our cohort with a p.Ser135Leu/p.*380Rext*50 genotype who has the highest CUMPDT120, despite an erythrocyte GALT activity below the limit of quantitation. This NBS-detected patient demonstrated only hyperbilirubinemia at diagnosis and currently demonstrates no long-term complications. The homozygous p.Ser135Leu patients in our cohort have a CUMPCDT120 in the control range without MDs, but both have a poor intellectual outcome. Their poor intellectual outcome may well be caused by their late initiation of dietary treatment. Since the early treated NBS variant patients and the p.Ser135Leu/ p.*380Rext*50 patient all have a CUMPCDT120 in the control range without demonstrating long-term complications, the question remains whether a CUMPCDT120 > 9 could be associated with a better clinical outcome if patients are treated early.

Limitations

This cohort includes 41 patients with varying ages and genotypes and the necessity to analyze results separately for adults and pediatric patients resulted in an even smaller sample size and hampered statistical power. To minimize the burdensome fasting and resting state in children in particular, and because previous research demonstrated a peak release of $^{13}CO_2$ into expired air at 120 minutes, we decided to measure the galactose oxidation capacity until 120 minutes[25]. After analyzing the results, a majority of the patients showed flattening in the area under the curve at 120 minutes, but not all patients had reached their maximum oxidation capacity at 120 minutes yet and this might have influenced the results.

The negative correlation we found between age and galactose oxidation capacity in patients with an identical genotype poses a limitation to the use of this test. Further studies are warranted to investigate whether this is truly due to a higher galactose oxidation capacity at a younger age.

Strengths

For the rarity of the disorder, we included a relatively large number of patients that represent the full genetic, biochemical, and clinical outcome spectrum of CG. The inclusion of multiple sibling pairs with a different intellectual outcome in some siblings, and the inclusion of twins provided more insight into whole body galactose oxidation. To ensure NBS variant and p.Ser135Leu patients did not influence the results, analyses were performed both with and without these patients.

Future perspectives

In our cohort, the galactose oxidation capacity of the classical patients was within a narrow

range and some differences between patients may be attributed to analytic variation. Also, sibling pairs with different clinical outcomes did not demonstrate significant differences in oxidation capacity. The limited number of patients may have affected the results or the 1-^{13}C breath test may not be sensitive enough to discriminate between classical patients.
Even though differences were small, the galactose oxidation capacity did differ between classical patients with identical genotypes. The use of universally 6-^{13}C labeled galactose might be able to differentiate within the group of classical patients by providing larger differences in $^{13}CO_2$ enrichment and thus enable better distinction between these patients.
In order to correctly interpret the results of the galactose oxidation breath test and to improve our understanding of galactose metabolism, the effects of age, body composition and genotype on galactose oxidation capacity should be further investigated in a larger cohort for which international cooperation will be necessary. Also, repeating the ^{13}C galactose breath test in the same patient cohort might provide valuable information with regard to the course of the galactose oxidation capacity with increasing age.
We demonstrated that NBS detects a group of variant patients with erythrocyte GALT activity up to 10%, with significantly higher galactose oxidation capacity than the classical patients, and no long-term complications so far. The question remains whether these patients are at risk for long-term complications or might even benefit from a less strict diet. Future studies addressing the galactose tolerance of these individuals are warranted to prevent overtreatment.

CONCLUSION

The ^{13}C galactose oxidation breath test is able to distinguish between classical patients, homozygous p.Ser135Leu patients and NBS variant patients, but is not able to differentiate between classical patients with poor and normal clinical outcomes in our cohort. Since NBS variant patients demonstrated galactose oxidation capacities in the control range, future studies are warranted to enable individualized prognostication and treatment, especially in these patients.

ACKNOWLEDGMENTS

The authors would like to thank Diana Ruffato Resende Campanholi for providing insight into the galactose oxidation calculations used in the study “Galactose oxidation using ^{13}C in healthy and galactosemic children”.

Funding
This study was supported by grants of Metakids and The AMC Foundation.

REFERENCES

1. Bosch AM. Classical galactosaemia revisited. *Journal of inherited metabolic disease*. 2006;29(4):516-525.
2. Fridovich-Keil JL, Walter JH. Part 7: Carbohydrates, Chapter 72: Galactosemia. The Online Metabolic and Molecular Bases of Inherited Disease, OMMBID; Valle D.L., Antonarakis S, Ballabio A, Beaudet A.L., Mitchell G.A.(Eds.). McGraw Hill, New York.
3. Hughes J, Ryan S, Lambert D, et al. Outcomes of siblings with classical galactosemia. *Journal of pediatrics*. 2009;154(5):721-726.
4. Kaufman FR, Kogut MD, Donnell GN, Goebelsmann U, March C, Koch R. Hypergonadotropic hypogonadism in female patients with galactosemia. *New England journal of medicine*. 1981;304(17):994-998.
5. Komrower GM, Lee DH. Long-term follow-up of galactosaemia. *Archives of disease in childhood*. 1970;45(241):367-373.
6. Kuiper A, Grunewald S, Murphy E, et al. Movement disorders and nonmotor neuropsychological symptoms in children and adults with classical galactosemia. *Journal of inherited metabolic disease*. 2019;42 (3):451-458.
7. Rubio-Gozalbo ME, Haskovic M, Bosch AM, et al. The natural history of classic galactosemia: lessons from the GalNet registry. *Orphanet journal of rare diseases*. 2019;14(1):86.
8. Welling L, Boelen A, Derks TG, et al. Nine years of newborn screening for classical galactosemia in The Netherlands: effectiveness of screening methods, and identification of patients with previously unreported phenotypes. *Molecular genetics and metabolism*. 2017;120(3):223-228.
9. Keevill NJ, Holton JB, Allen JT. The investigation of UDP glucose and UDP-galactose concentration in red blood cells of patients with classical galactosaemia. *Clinica chimica acta*. 1993;221 (1–2):135-142.
10. Slepak TI, Tang M, Slepak VZ, Lai K. Involvement of endoplasmic reticulum stress in a novel classic Galactosemia model. *Molecular genetics and metabolism*. 2007;92(1–2):78-87.
11. Coss KP, Hawkes CP, Adamczyk B, et al. *N*-glycan abnormalities in children with galactosemia. *Journal of proteome research*. 2014;13(2): 385-394.
12. Lai K, Elsas LJ, Wierenga KJ. Galactose toxicity in animals. *International union of biochemistry and molecular biology (IUBMB) Life.* 2009;61(11):1063-1074.
13. Yuzyuk T, Viau K, Andrews A, Pasquali M, Longo N. Biochemical changes and clinical outcomes in 34 patients with classic galactosemia. *Journal of inherited metabolic disease*. 2018;41(2):197-208.
14. Barbouth DS, Velazquez DL, Konopka S, Wilkinson JJ, Carver VH, Elsas LJ. Screening newborns for galactosemia using total body galactose oxidation to CO2 in expired air. *Pediatric research*. 2007;62(6):720-724.
15. Berry GT, Singh RH, Mazur AT, et al. Galactose breath testing distinguishes variant and severe galactose-1-phosphate uridyltransferase genotypes. *Pediatric research*. 2000;48(3):323-328.
16. Resende-Campanholi DR, Porta G, Ferrioli E, Pfrimer K, Ciampo LA, Junior JS. Galactose oxidation using (13)C in healthy and galactosemic children. *Brazilian journal of medical and biological research*. 2015; 48(3):280-285.

17. Webb AL, Singh RH, Kennedy MJ, Elsas LJ. Verbal dyspraxia and galactosemia. *Pediatric research*. 2003;53(3):396-402.
18. Lai K, Langley SD, Singh RH, Dembure PP, Hjelm LN, Elsas LJ2nd. A prevalent mutation for galactosemia among black Americans. *Journal of Pediatrics*. 1996;128(1):89-95.
19. Morrison DJ, Dodson B, Slater C, Preston T. (13)C natural abundance in the British diet: implications for (13)C breath tests. *Rapid Communications in Mass Spectrometry*. 2000;14(15):1321-1324.
20. Welling L, Bernstein LE, Berry GT, et al. International clinical guideline for the management of classical galactosemia: diagnosis, treatment, and follow-up. *Journal of inherited metabolic disease*. 2017;40(2): 171-176.
21. Reckman G, Koehorst M, Priebe M, Schierbeek H, Vonk R.13C protein oxidation in breath: is it relevant for the whole body protein status? *Journal of biomedical science and engineering*. 2016;9:160-169.
22. van der Schoor SR, de Koning BA, Wattimena DL, Tibboel D, van Goudoever JB. Validation of the direct nasopharyngeal sampling method for collection of expired air in preterm neonates. *Pediatric research*. 2004;55(1):50-54.
23. Elsas LJ, Ellerine NP, Klein PD. Practical methods to estimate whole body leucine oxidation in maple syrup urine disease. *Pediatric research*. 1993;33(5):445-451.
24. David C, Coleman BF. Chapter 14: tracer studies with 13C enriched substrates by Thomas W. Boutton. In: Coleman DC, ed. Carbon Isotope Techniques. San Diego, CA: Academic Press; 1991.
25. Berry GT, Nissim I, Mazur AT, et al. In vivo oxidation of [13C] galactose in patients with galactose-1-phosphate uridyltransferase deficiency. *Biochemical and molecular medicine*. 1995;56(2): 158-165.
26. Coelho AI, Bierau J, Lindhout M, Achten J, Kramer BW, Rubio-Gozalbo ME. Classic Galactosemia: study on the late prenatal development of GALT specific activity in a sheep model. *The anatomical record: advances in integrative anatomy and evolutionary biology*. 2017;300(9):1570-1575.

SUPPLEMENTARY MATERIAL

Supplementary Table 1. CUMPDT120 and clinical outcomes in siblings

Patient group:	Classical	Variants	Classical	Classical	Classical	Classical	Classical
Sibling 1	Pre-NBS	NBS	Pre-NBS	Pre-NBS	FS	Pre-NBS	Pre-NBS
Sibling 2	Pre-NBS	NBS	FS	FS	FS	FS	FS
Age at diagnosis (days)							
Sibling 1	24	9	14	-	0	7	14
Sibling 2	24	9	0	0	0	0	0
CUMPCDT120							
Sibling 1	1.04	14.87	0.67	0.17	0.119	0.30	0.50
Sibling 2	2.02	14.61	0.53	0.21	0.121	0.43	0.29
Intellectual outcome							
Sibling 1	Poor	Normal	Normal	Poor	Poor	Normal	Normal
Sibling 2	Normal	Normal	Poor	Poor	Poor	Poor	Normal
MD							
Sibling 1	Yes	No	Yes	Yes	Yes	Yes	No
Sibling 2	Yes	No	Yes	Yes	Yes	Yes	-

Notes. NBS: Newborn screening, FS: Family screening, CUMPCD: Cumulative Percentage Dose Recovered, Poor: IQ<85, Normal: IQ≥85, MD: Movement Disorder, - : missing data.

Chapter 7

Gray and white matter are both affected in classical galactosemia: an explorative study on the association between neuroimaging and clinical outcome

M.M. Welsink-Karssies, A. Schrantee, M.W.A. Caan, C.E.M. Hollak, M.C.H. Janssen, E. Oussoren, M.C. de Vries, S.D. Roosendaal, M. Engelen*, A.M. Bosch*

* Both authors contributed equally

In preparation

A modified version of this manuscript has been accepted for publication in the journal Molecular Genetics and Metabolism

ABSTRACT

Classical Galactosemia (CG) is an inherited disorder of galactose metabolism caused by a deficiency of the galactose-1-phosphate uridylyltransferase (GALT) enzyme resulting in neurocognitive complications. As in many Inborn Errors of Metabolism, the metabolic pathway of CG is well-defined, but the pathophysiology and high variability in clinical outcome are poorly understood. The aim of this study was to investigate structural changes of the brain of CG patients on MRI and their association with clinical outcome. In this prospective cohort study an MRI protocol was developed to evaluate gray matter (GM) and white matter (WM) volume of the cerebrum and cerebellum, the volume of WM hyperintensities, WM microstructure and myelin content with the use of conventional MRI techniques, diffusion tensor imaging (DTI) and quantitative T1 mapping. The association between several neuroimaging parameters and both neurological and intellectual outcome was investigated.

Twenty-one patients with CG (median age 22 years, range 8-47) and 24 controls (median age 30, range 16-52) were included. Compared with controls, the WM volume of CG patients was lower, the microstructure of WM was impaired both in the whole brain and corticospinal tract (CST) and the lower R1 values of WM, GM and the CST were indicative of less myelin. The volume of WM hyperintensities were comparable between patients and controls. The 9 out of 16 patients with a poor neurological outcome (defined as the presence of a tremor and/or dystonia) demonstrated a lower WM volume, an impaired WM microstructure and lower R1 values of the WM compared with 7 out of 16 patients without movement disorders. In 15 out of 21 patients with a poor intellectual outcome (defined as an IQ<85) both GM and WM were affected with a lower cerebral and cerebellar WM and GM volume compared with 6 out of 21 patients with an IQ≥85. Both the severity of the tremor (as indicated by the Tremor Rating Scale) and IQ (as continuous measure) were associated with several neuroimaging parameters such as GM volume, WM volume, CSF volume, WM microstructure parameters and R1 values of GM and WM.

In this explorative study performed in patients with CG, not only WM but also GM pathology was found, with more severe brain abnormalities on MRI in patients with a poor neurological and intellectual outcome. The finding that structural changes of the brain were associated with the severity of long-term complications indicates that quantitative MRI techniques could be of use to explain neurological and cognitive dysfunction as part of the disease spectrum. Based on the clinical outcome of patients, the absence of widespread WM hyperintensities and the finding that both GM and WM are affected, CG could be is primarily a GM disease with secondary damage to the WM as a result of neuronal degeneration. To investigate this further the course of GM and WM should be evaluated in longitudinal research, which could also clarify if CG is a neurodegenerative disease.

INTRODUCTION

Classical Galactosemia (CG, OMIM 230400) is an inherited disorder of galactose metabolism caused by pathogenic variations in the *GALT* gene leading to a severe deficiency of the enzyme galactose-1-phosphate uridylyltransferase (GALT, EC 2.7.7.12). An early initiation of a galactose-restricted diet prevents severe illness in the neonatal period, but does not prevent long-term complications such as speech- and language delay, movement disorders and cognitive impairment[1-4]. In CG, the central nervous system is primarily affected and primary ovarian insufficiency is frequently reported in females, implicating organ specific tissue damage. The clinical outcome of patients ranges from normal to severely impaired, even within families with identical *GALT* gene variations, which is poorly understood[2,5]. Presumably, both the persistent elevation of harmful metabolites such as galactose-1-phosphate (Gal-1-P) and the reduced production of important substrates downstream of the GALT deficiency such as UDP sugars, contribute to the long-term complications observed in CG patients[5-8]. Both mechanisms may affect the galactosylation of proteins and lipids[5,9] and galactosylation abnormalities have been reported in CG patients[10,11].

Recently it has been suggested that the extent of the galactosylation abnormalities may correlate with the clinical outcome of patients[12,13]. A number of studies investigated the brain of CG patients with the use of MRI and found abnormalities, but the frequency of abnormalities varied between studies and most studies included a limited number of patients[2,14-18]. The two largest cohorts found an abnormal cerebral white matter signal intensity on T1-weighted and T2-weighted scans in 79-83% of the patients and interpreted this finding as abnormal (delayed/ deficient) myelination. Other observed abnormalities included cerebral atrophy in 16-33%, cerebellar atrophy in 12-58% and white matter hyperintensities in 16% of the patients[2,15]. The previously mentioned studies all used a qualitative approach and described the frequency of brain abnormalities on MRI. These studies demonstrated that visible white matter hyperintensities on MRI occur infrequently (16%) but subtle signal intensity abnormalities were frequently observed. To investigate this further, more quantitative measures derived from MRI scans that are able to assess subtle differences are highly relevant for studying the brain in CG patients. The studies that have used quantitative MRI techniques, such as voxel based morphometry, revealed white matter pathology that correlated with (cognitive) outcome[19], and gray matter abnormalities, more specifically gray matter density decreases as well as increases in CG patients compared to controls[20]. Until now it is unclear whether white matter is predominantly affected in CG or if gray matter is primarily affected, with secondary changes in the white matter due to axonal deterioration. As brain abnormalities on MRI have been observed both in patients with and without long-term complications[18,21], the question remains whether structural changes on MRI are associated with the clinical outcome of patients and are able to explain the variability

in clinical outcome.
As pathological processes may be present in normal appearing white- and gray matter, the aim of this study is to gain knowledge on the pathophysiology of CG by assessing both white matter and gray matter of CG patients with the use of quantitative neuroimaging parameters, and to investigate the association between structural changes on MRI and clinical outcome.

METHODS

Study recruitment and ethics

For this study, CG patients who visited the multidisciplinary galactosemia outpatient clinic in the Amsterdam University Medical Centers for research or clinical purposes were recruited between April 2018 and November 2018. All included patients (and/or parents) gave informed consent before study enrollment and consented with the use of their clinical data for research purposes. Healthy controls included in a study on adrenoleukodystrophy[22] were used as control group in this study as the same MRI protocol was used. These male controls were recruited via advertisements, received a structured history and physical examination and were included if neurological co-morbidity was absent. The healthy female control included in this study underwent the MRI protocol for testing purposes and was included as well. This study was approved by the local medical ethics committee.

Inclusion and exclusion criteria

Galactosemia patients with GALT enzyme activity in erythrocytes <15% of the reference mean and/or two known pathogenic variations in the *GALT* gene, aged eight years and older and able to undergo an MRI scan of the brain were eligible for participation in this study. All included patients had received intellectual and/or neurological testing. Patients with a second diagnosis influencing clinical outcome, with anxiety or claustrophobia or with MRI contraindications, such as dental braces were not eligible to participate.

Clinical outcome

In order to study differences in clinical outcome, patients were divided into groups based on their intellectual and neurological outcome. In patients who received a standardized age-specific intelligence test, the IQ was used as derivative of intelligence. The used intelligence tests were the Wechsler Preschool and Primary Scale of Intelligence, the Wechsler Intelligence Scale for Children and the Wechsler Adult Intelligence Scale. A poor intellectual outcome was defined as an IQ<85 and a normal intellectual outcome as an IQ≥85. The presence or absence of movement disorders was used as a determinant for neurological outcome. Information was retrieved from the medical records of the

included patients. Patients were investigated by one (pediatric) neurologist and in addition to the neurological examination a standardized tremor rating scale (TRS), the Fahn-Tolosa-Marin Clinical Rating Scale for Tremors[23] and dystonia rating scale (DRS), the Fahn-Marsden Rating Scale[24] was performed.

MRI acquisition

Protocol

A 3.0 Tesla MRI scanner (Philips Ingenia, Philips, Best, the Netherlands) with a 32-channel head coil was used to perform MR imaging of the brain. The following MRI sequences were included in the protocol: 1) a 3D T1 weighted anatomical scan (TR/TE=7.5/3.4ms, flip angle=8°, in plane resolution=1.06 x 1.06 mm; slice thickness=1.1 mm), 2) a 3D fluid attenuated inversion recovery (FLAIR) scan (TR/TE=4800/355 ms; TI=1650 ms; 160 sagittal sections; in-plane resolution=1.1 × 1.1 mm, slice thickness=1.12 mm), 3) a single-shot echo planar imaging Diffusion Tensor Imaging (DTI) scan (TR/TE=6285/82ms; 2.5 mm^3 isotropic resolution, 32 gradient directions, number of b-factors 2; b-value 1000 s/mm^2) and 4) T1 mapping using a variable flip angle protocol (TR/TE=23.3/2.3ms; flip angle=7° and 20°; 1.1 mm^3 isotropic resolution. Additional B1-mapping consisting of a sagittal DREAM sequence was needed for quantification.

Volumetric analysis of T1-weighted MR images

The T1-weighted images were used to measure brain tissue volume. T1-weighted images were segmented into white matter (WM), gray matter (GM) and cerebral spinal fluid (CSF) using the Computational Anatomy Toolbox (CAT 12, SPM). Volumetric measurements were performed after lesion masking of the previously defined and marked WM hyperintensities. Cerebellar cortical and WM volumes were extracted from Freesurfer (v5.3.0) segmentation. The CAT12 quality control was evaluated which included resolution, noise and bias. An image quality (%bias) ≥75% equals satisfactory results. Also, the GM/WM borders on the coronal and transversal T1 images were visually inspected.

DTI analysis: microstructure of the whole brain and corticospinal tract

The diffusion weighted images were denoised with the use of MRtrix3[25] and corrected for distortions caused by eddy currents and head motion using FSL[26]. Outliers as defined by head movement equal or more than three standard deviations, were replaced by Gaussian Process predicted values[27]. Using an in-house Matlab script b0 ringing artefacts were corrected and tensors were fitted to the diffusion data to create fractional anisotropy (FA), mean diffusivity (MD), axial diffusivity (AD) and radial diffusivity (RD) images (FDT - FMRIB's Diffusion Toolbox, FSL). (S)QUAD was used for quality control and to identify potential outliers[28]. The sum of squared error (SSE) values were

used to assess the quality of the tensor fit. SSE values were considered outliers in case of a value ≥p95.
Preprocessed data were analyzed using Tract-Based Spatial Statistics (TBSS)[29], part of FSL[30]. The FA maps of all participants were aligned to the MNI152 standard space using non-linear registration, such that every image was aligned to every other one, using the most representative study-specific FA image as a target. A mean FA skeleton was retrieved and the FA images of all participants were projected onto the mean FA skeleton to create skeletonized FA data for each participant. We applied the non-linear FA warping and skeletonized FA projection to MD, RD and AD images resulting in skeletonized data of all these measures. Subsequently, a cortical spinal tract region of interest (ROI) mask was obtained from the JHU ICBM-DTI-81 white-matter labels atlas[31-33] and ROI data for all subjects were extracted and fed into statistical analysis.

Quantitative T1 mapping
Quantitative T1 mapping (qT1) of GM in the whole brain and WM in the whole brain and corticospinal tract (CST) was performed. Previous research has demonstrated that up to 90% of the R1 signal (1/T1) is explained by variations in myelin content[34]. T1-maps were computed using the fitMRI toolbox[35], in Matlab (The MathWorks Inc., Natick, MA, USA) using SPM8 (www.fil.ion.ucl.ac.uk/spm). Maximum-likelihood estimation of T1 was performed from the co-registered series of two multi-echo spoiled gradient recalled echo (SPGR) images. Flip angles were scaled per voxel using the local percentage of the nominal flip angle as obtained from the B1-map[36]. Mean R1 values were extracted from a GM and WM mask as well as from the CST ROIs. To ensure the quality of the data, the R1 scans were visually inspected.

Identification of white matter hyperintensities
To identify and quantify WM hyperintensities (deep WM and periventricular WM) on FLAIR images, the lesion prediction algorithm of the lesion segmentation toolbox (LST, SPM12) was used and adjusted using manual correction. The manual correction was performed by one investigator (MWK) and were visually inspected by a neuro-radiologist (SR). Hereafter, the WM hyper intensity volumes of both deep WM and periventricular WM were re-calculated and a 15μL filter was applied to identify potential relevant WM hyperintensities.

Statistical analysis
SPSS version 26 (SPSS Inc. Chicago, Illinois, USA) was used to perform all statistical analyses. Figures were generated using Graphpad Prism (Version 5, Graphpad Software Inc., La Jolla, California, USA). Before statistical analyses of the data, we reviewed the available literature for normative data on WM, GM, cerebellar and WM hyper intensity volumes and DTI measures (FA, MD, RD, AD) of healthy subjects considering age

and gender are possible confounders. We present the normative data retrieved from the literature as a red line in the graphical presentation of the data. The descriptive statistics of patient and control characteristics were reported. Non-parametric tests were used to test for differences between (sub)groups and the Spearman's rank coefficient test was used to test for correlations as the sample size was small and most data followed a non-normal distribution. All test presumed a two-tailed probability and a p-value < 0.05 was considered a significant difference. First, we investigated the following neuroimaging parameters in patients and controls; white matter volume (WMV), gray matter volume (GMV), cerebral spinal fluid volume (CSFV) and cerebellar volumes (cerebellar cortex volume; CCV and cerebellar white matter volume; CWMV), the volume of WM hyperintensities, DTI measures and R1 values. Second, we investigated these neuroimaging parameters in patients with and without movement disorders. The severity of the movement disorder as measured by the tremor- and dystonia rating scale were evaluated as well. Third, we investigated the neuroimaging parameters in patients with a poor (IQ<85) and normal intellectual outcome (IQ≥85) and intellectual outcome with IQ as continuous outcome measure.

RESULTS

A total of 21 CG patients (aged 8-47 years), 9 males and 12 females were included. The matching of patients and controls based on a maximum age difference of five years resulted in a control group (aged 16-52 years) which included 23 males and 1 female. Demographics are shown in **Table 1**.

Patients versus controls

First, we evaluated differences between patients and controls (**Table 1**). Both age and gender were not equally distributed between patients and controls.

Volumetric measurements

The volumetric analyses of WM, GM, CSF, cerebellar WM (CWM) and cerebellar cortex (CC) are shown in **Table 1**. As controls demonstrated significantly larger total intracranial volumes (TIV) than patients, both absolute volumes and fractional controlling volumes (%WM, %GM and %CSF) are reported.

The controls demonstrated similar patterns in white matter volume (WMV) and gray matter volume (GMV) as described in the literature[37](**Figure 1**). In patients, the absolute WMV and %WMV was significantly lower when compared with controls (p<0.0005). The course of %WMV of patients with different ages was comparable to the pattern described in the literature and as demonstrated for controls, but was overall lower (**Figure 1A**). The %GMV, but not the absolute GMV was significantly higher compared with

controls (p<0.0005). The decline in %GMV followed the same pattern as described in the literature and as demonstrated for controls (**Figure 1B**), but revealed a steeper decline for male patients when compared with male controls and female patients (**Figure 1C**). The cerebellar cortex volume (CCV) of patients was significantly lower when compared with controls (p=0.004) (**Table 1**). Total intracranial volume (TIV) was not correlated with age, but males did demonstrate significantly larger TIVs compared with females (1524 cm^3 versus 1347 cm^3). As gender was significantly different between patients and controls and only one female control was included, additional analyses were performed for males separately. Male patients had significantly lower absolute WMV and %WMV (p<0.0005), significantly higher absolute GMV (p=0.038) and %GMV (p<0.0005), and significantly lower absolute CSFV (p=0.0003) and %CSFV (p=0.034) when compared with controls. As age is a confounder for volumetric measurements and the male patients with a median age of 15 years (range 8-31) were significantly younger (p<0.0005) than the male controls with a median age of 30 years (range 16-52), patients were age and gender matched, resulting in five male patients and five male controls (age matched within 2 years) (**Supplementary Table 1**). In this group of matched males, the %WMV remained significantly lower in patients when compared with controls (p=0.028). All other volumetric measurements were not significantly different. The %GMV in patients seems to demonstrate a steeper decline when compared with controls.

White matter microstructure in DTI

The FA of the whole brain of patients was significantly lower when compared with controls (p=0.002) (**Table 1**, **Supplementary Figure 1A**). The FA of the corticospinal tract (CST) was significantly lower (p=0.012) (**Supplementary Figure 1B**) and MD (p=0.004) and RD (p=0.009) significantly higher in patients when compared with controls.

QT1: myelin mapping - R1

The R1 values of WM, GM and CST of patients were significantly lower when compared with controls (p<0.027) (**Table 1**). As both age and gender were significantly different between patients and controls, analyses were repeated for males after which the significant difference in WM and GM remained (p<0.019). As age remained significantly different, the R1 of WM in patients with different ages was demonstrated in **Supplementary Figure 1C**.

Figure 1A

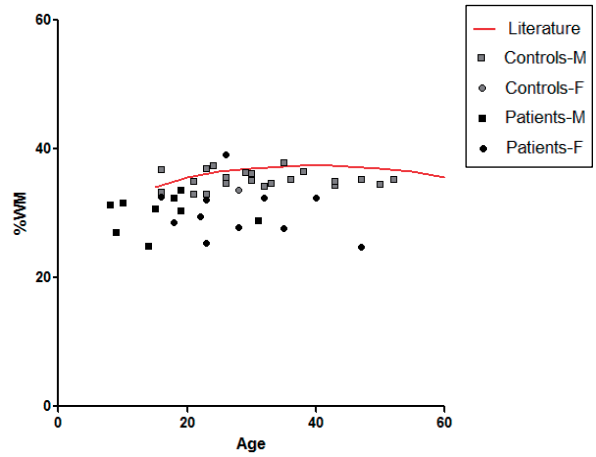


Figure 1B

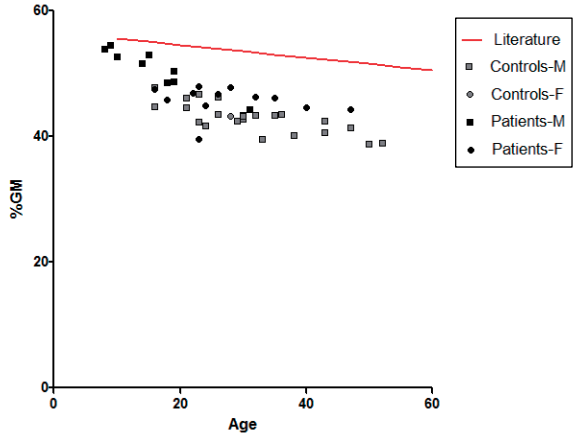


Figure 1C

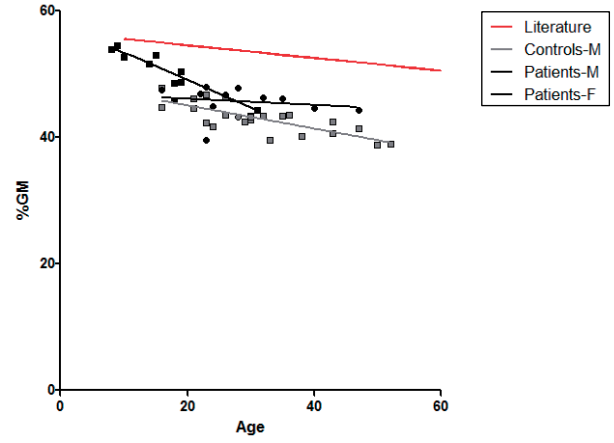


Figure 1. Volumetric measurements of CG patients (M: males and F: females), controls (M: males and F: females) and data from the literature based on healthy subjects (red line)[37].

Table 1. Patients versus controls

	Patients	Controls	P-value
n	21	24	-
Gender			**< 0.0005**
Male, *n* (%)	9 (43)	23 (96)	
Age in years	22 (8 – 47)	30 (16 – 52)	**0.006**
Volumetric measurements			
- TIV in cm^3	1355 (1176 – 1644)	1554 (1325 – 1776)	**< 0.0005**
- WMV	431 (299 – 533)	541 (467 – 654)	**<0.0005**
- GMV	636 (465 – 801)	668 (548 – 770)	0.426
- CSFV	301 (183 – 410)	330 (259 – 441)	**0.016**
- %WMV	30.7 (24.8 – 39.0)	35.2 (32.9 – 37.9)	**< 0.0005**
- %GMV	47.5 (39.5 – 54.4)	43.2 (38.8 – 47.7)	**< 0.0005**
- %CSFV	21.5 (14.3 – 34.9)	21.7 (18.3 – 26.6)	0.900
- CWMV in mL	29.1 (20.5 – 46.0)	31.7 (26.3 – 38.2)	0.172
- CCV in mL	109.7 (81.7 – 132.0)	117.9 (97.0 – 139.6)	**0.004**
DTI measures			
- FA whole brain	0.380 (0.335 – 0.416)	0.396 (0.376 – 0.414)	**0.002**
- MD whole brain*	0.849 (0.807 – 0.935)	0.849 (0.818 – 0.886)	0.554
- RD whole brain*	0.671 (0.610 – 0.752)	0.653 (0.619 – 0.694)	0.088
- AD whole brain*	1.227 (1.169 – 1.300)	1.241 (1.211 – 1.271)	0.062
- FA CST	0.541 (0.508 – 0.601)	0.560 (0.532 – 0.600)	**0.012**
- MD CST*	0.803 (0.760 – 0.848)	0.779 (0.760 – 0.820)	**0.004**
- RD CST*	0.534 (0.466 – 0.576)	0.500 (0.462 – 0.545)	**0.009**
- AD CST*	1.364 (1.295 – 1.407)	1.348 (1.292 – 1.389)	0.152
n	21	20	-
R1 values			
- WM	0.892 (0.692 – 1.007)	0.995 (0.807 – 1.153)	**< 0.0005**
- GM	0.629 (0.475 – 0.693)	0.674 (0.532 – 0.785)	**0.001**
- CST	0.897 (0.649 – 1.096)	0.925 (0.771 – 1.105)	**0.027**

Notes. Data reported in median (range). TIV: Total intracranial volume, WMV: absolute white matter volume, GMV: absolute gray matter volume, CSF: absolute cerebral spinal fluid volume, %WMV: relative white matter volume, %GMV: relative gray matter volume, %CSF: relative cerebral spinal fluid volume, CWMV: cerebellar white matter volume, CCV: cerebellar cortex volume, FA: fractional anisotropy, MD: mean diffusivity, RD: radial diffusivity, AD: axial diffusivity, CST: corticospinal tract. *MD, RD and AD values x 10^{-3}, R1 (1/T1) in seconds, WM: white matter, GM: gray matter.

Neurological outcome

Second, we evaluated differences between patients with and without movement disorders (**Table 2**). A movement disorder was observed in 9 out of 16 patients and varied from a mild tremor and dystonia to a severe tremor and dystonia interfering with daily functioning. Age and gender were equally distributed between patients with and without movement disorders (**Table 2**).

Volumetric measurements

Patients with a movement disorder had significantly lower absolute WMV (p=0.005) and %WMV (p=0.003) compared with patients without a movement disorder. The trend in %WMV was overall lower for patients with movement disorders (**Figure 2A**). The tremor rating scale was negatively correlated with both absolute WMV and %WMV (p<0.0005) (**Figure 2B**) and CWMV (p=0.014) (**Figure 2C**) and positively

correlated with %CSFV (p=0.037). The dystonia rating scale was negatively correlated with %WMV (p=0.044) and was positively correlated with absolute CSFV (p=0.029).

Figure 2A

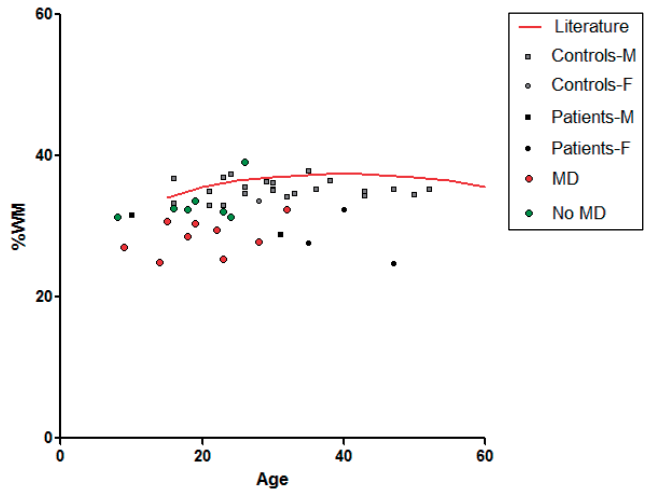


Figure 2A. Sixteen out of 21 patients received a neurological examination, 9/16 had a movement disorder (MD) and 7/16 had no movement disorder (no MD). Data from the literature based on healthy subjects (red line)[37].

Figure 2B

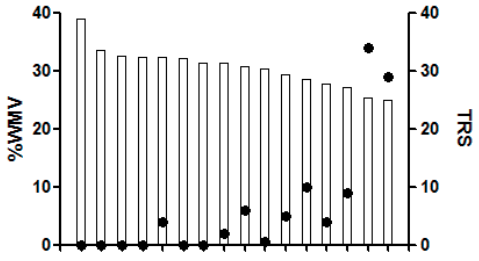


Figure 2C

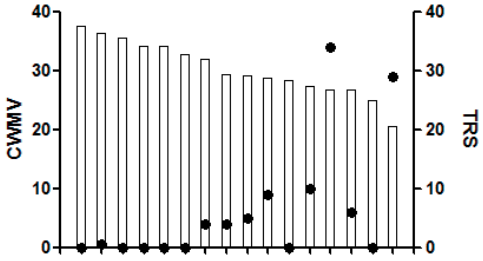


Figure 2B/2C. Every bar represents the %white matter volume (%WMV)/cerebral white matter volume (CWMV) of one patient on the left Y- axis and the corresponding Tremor Rating Scale (TRS) on the right Y-axis.

White matter microstructure in DTI

In patients with movement disorders, the FA whole brain was significantly lower when compared with patients without movement disorders (**Supplementary Figure 2**). There was a negative correlation between the tremor rating scale (TRS) and FA whole brain

(p=0.001) and FA CST (p=0.020), and a positive correlation with RD whole brain (p=0.007) and RD CST (p=0.026). The DTI measures of the whole brain and CST were not correlated with the dystonia rating scale.

QT1: myelin mapping - R1

In patients with movement disorders, the R1 of WM was significantly lower when compared with patients without movement disorders (**Table 2**). The GM was higher in patients without movement disorders, but the difference was not significant (p=0.050). There was a negative correlation between the tremor rating scale (TRS) and GM R1 (p=0.018) and WM R1 (p=0.011).

Table 2. Patients with versus patients without movement disorders

	Patients with movement disorders	Patients without movement disorders	P-value
n	9	7	-
Gender			
Male, *n* (%)	4 (44)	3 (43)	p>0.999
Age in years	19 (9 – 32)	19 (8 – 26)	p>0.999
Volumetric measurements			
- TIV in cm^3	1304 (1176 – 1591)	1442 (1284 – 1644)	0.064
- WMV	396 (299 – 484)	466 (431 – 533)	**0.005**
- GMV	635 (465 – 801)	667 (599 – 798)	0.266
- CSFV	301 (212 – 410)	268 (183 – 351)	0.223
- %WMV	28.5 (24.9 – 32.3)	32.4 (31.3 – 39.0)	**0.003**
- %GMV	47.7 (39.5 – 54.4)	47.9 (44.9 – 53.8)	0.958
- %CSFV	23.3 (16.3 – 34.9)	19.0 (14.3 – 23.6)	0.064
- CWMV in mL	28.8 (20.5 – 36.4)	34.2 (24.9 – 37.6)	0.153
- CCV in mL	108.1 (81.7 – 115.4)	118.9 (85.5 – 132.0)	0.125
DTI measures			
- FA whole brain	0.375 (0.335 – 0.396)	0.399 (0.362 – 0.416)	**0.039**
- MD whole brain*	0.847 (0.833 – 0.902)	0.841 (0.807 – 0.866)	0.368
- RD whole brain*	0.671 (0.641 – 0.733)	0.644 (0.610 – 0.671)	0.081
- AD whole brain*	1.221 (1.169 – 1.242)	1.226 (1.198 – 1.256)	0.491
- FA CST	0.546 (0.508 – 0.590)	0.566 (0.517 – 0.601)	0.315
- MD CST*	0.799 (0.776 – 0.841)	0.773 (0.760 – 0.821)	0.266
- RD CST*	0.522 (0.480 – 0.576)	0.497 (0.466 – 0.557)	0.266
- AD CST*	1.368 (1.311 – 1.392)	1.351 (1.326 – 1.371)	0.634
R1 values			
- WM	0.860 (0.721 – 0.911)	0.924 (0.692 – 1.007)	**0.039**
- GM	0.619 (0.543 – 0.645)	0.652 (0.475 – 0.693)	0.050
- CST	0.837 (0.795 – 0.950)	0.902 (0.649 – 1.096)	0.125

Notes. Data reported in median (range). TIV: Total intracranial volume, WMV: absolute white matter volume, GMV: absolute gray matter volume, CSF: absolute cerebral spinal fluid volume, %WMV: relative white matter volume, %GMV: relative gray matter volume, %CSF: relative cerebral spinal fluid volume, CWMV: cerebellar white matter volume, CCV: cerebellar cortex volume, FA: fractional anisotropy, MD: mean diffusivity, RD: radial diffusivity, AD: axial diffusivity, CST: corticospinal tract. *MD, RD and AD values x 10^{-3}, R1 (1/T1) in seconds, WM: white matter, GM: gray matter.

Intellectual outcome

Third, we investigated differences between patients with a poor (IQ<85) and normal intellectual outcome (IQ≥85) (**Table 3**). In total, 6 out of 21 patients had a normal intellectual outcome. Age and gender were not equally distributed between patients with an IQ≥85 and an IQ<85 (**Table 3**).

Volumetric measurements

The results presented in **Figure 3A** demonstrate the non-linear, quadratic trend in %WMV, but the overall trend seems to be lower for patients with an IQ<85 than for patients with an IQ>85. Patients with an IQ<85 demonstrated a significantly lower absolute GMV (**Figure 3B**), absolute WMV, %GMV, CWMV and CCV and a significantly higher %CSFV when compared with patients with an IQ≥85 (p<0.043). There was a positive correlation between IQ (as continuous measure) and absolute WMV, absolute GMV, %WMV (**Figure 3C**), CCV (**Figure 3D**) and CWMV (p<0.027), and a negative correlation between IQ and %CSFV (p=0.015). As the male patients were equally distributed with regard to age (p>0.999) analyses were repeated for this group which demonstrated a significantly lower absolute WMV, absolute GMV, %WMV (**Supplementary Figure 3**), CWMV and CCV in male patients with an IQ<85 when compared with male patients with an IQ>85 (p<0.027). The positive correlation between the IQ and absolute WMV, absolute GMV, %WMV, CWMV and CCV remained (p<0.020).

White matter microstructure in DTI

The DTI measures (FA, MD and RD) of the whole brain were not significantly different between patients with an IQ<85 and an IQ>85 and were not correlated with IQ as continuous measure.

QT1: myelin mapping - R1

The R1 values were not significantly different between patients with an IQ<85 and an IQ>85 (p>0.119). IQ as continuous measure was positively correlated with WM (p=0.044) and GM (p=0.014).

Table 3. Patients with a poor versus patients with a normal intellectual outcome

	Patients with IQ<85	Patients with IQ≥85	P-value
n	15	6	-
IQ	71 (49-83)	92 (88-98)	-
Gender			
Male, *n* (%)	4 (27)	5 (83)	**0.046**
Age in years	24 (9 – 47)	18.5 (8 – 22)	0.056
Volumetric measurements			
- TIV in cm^3	1304 (1176 – 1487)	1502 (1355 – 1644)	**0.004**
- WMV	400 (299 – 501)	476 (399 – 533)	**0.013**
- GMV	622 (465 – 691)	778 (635 – 801)	**0.002**
- CSFV	301 (183 – 410)	286 (214 – 320)	0.392
- %WMV	28.8 (24.8 – 39.0)	31.4 (29.4 – 33.6)	0.213
- %GMV	46.2 (39.5 – 54.4)	49.5 (46.9 – 53.8)	**0.036**
- %CSFV	23.3 (14.3 – 34.9)	18.3 (14.9 – 23.7)	**0.043**
- CWMV in mL	27.7 (20.5 – 46.0)	35.3 (28.3 – 38.7)	**0.036**
- CCV in mL	99.7 (81.7 – 118.9)	120.1 (113.8 – 132.0)	**0.001**
DTI measures			
- FA whole brain	0.380 (0.335 – 0.416)	0.379 (0.362 – 0.405)	0.586
- MD whole brain*	0.849 (0.807 – 0.935)	0.848 (0.841 – 0.881)	0.938
- RD whole brain*	0.671 (0.610 – 0.752)	0.665 (0.642 – 0.693)	0.697
- AD whole brain*	1.227 (1.169 – 1.300)	1.230 (1.126 – 1.259)	0.815
R1 values			
- WM	0.873 (0.692 – 1.007)	0.906 (0.872 – 0.954)	0.119
- GM	0.625 (0.475 – 0.693)	0.647 (0.624 – 0.659)	0.119
- CST	0.852 (0.649 – 0.990)	0.905 (0.856 – 1.096)	0.276

Notes. Data reported in median (range). IQ: intellectual quotient, TIV: Total intracranial volume, WMV: absolute white matter volume, GMV: absolute gray matter volume, CSF: absolute cerebral spinal fluid volume, %WMV: relative white matter volume, %GMV: relative gray matter volume, %CSF: relative cerebral spinal fluid volume, CWMV: cerebellar white matter volume, CCV: cerebellar cortex volume, FA: fractional anisotropy, MD: mean diffusivity, RD: radial diffusivity, AD: axial diffusivity, CST: corticospinal tract. * MD, RD and AD values x 10^{-3}, R1 (1/T1) in seconds, WM: white matter, GM: gray matter.

White matter hyperintensities

In total, 8 patients (38%) and 5 controls (21%) had deep white matter hyperintensities (DWMH) and 15 patients (71%) and 19 controls (79%) had periventricular white matter hyperintensities (PVWMH). The DWMH and PVWMH volumes were higher in patients compared with controls, but not significantly different (**Supplementary Table 2.1** and **2.2**). The DWMH and PVWMH volumes were not significantly different between patients with and without movement disorders nor between patients with an IQ<85 and IQ>85. IQ (as continuous measure) was not correlated with DWMH and PVWMH volumes.

Quality control

For the volumetric analyses, the program CAT12 reports the image and preprocessing quality as a percentage from 0-100 with regard to resolution, noise and bias. Patients had significantly more bias than controls. The %bias was positively correlated with age and %WM and negatively correlated with %GM. Considering an image quality (%bias) of

≥75% equals satisfactory results and participants with a %bias <75% demonstrated more artefacts upon visual inspection, a %bias of ≥75% was used as cut-off value which led to the exclusion of seven patients. The MRI of the youngest patient in our cohort showed motion artefacts and the image quality was just below the cut-off of 75%. All other excluded patients had a poor intellectual outcome and of the 4/6 patients that received a neurological examination all four had a movement disorder. After exclusion, the significant differences between patients and controls, and between patients with and without movement disorders remained, but the correlation between IQ and %WM and CWMV, and between TRS and CWMV were no longer significant. The visual inspection of the T1-weighted images revealed an abnormal white matter intensity signal in patients compared with controls, which was more prevalent in patients with a poor clinical outcome (**Figure 4**). For the DTI measures, the SSE values for patients and controls were not significantly different. For the youngest patient in our cohort, both the whole brain and the CST SSE values were clear outliers. The exclusion of this patient did not change any of the results. For the T1 mapping, the R1 scan of the youngest patient in our cohort demonstrated motion artefacts. The exclusion of this patient did not change any of the results.

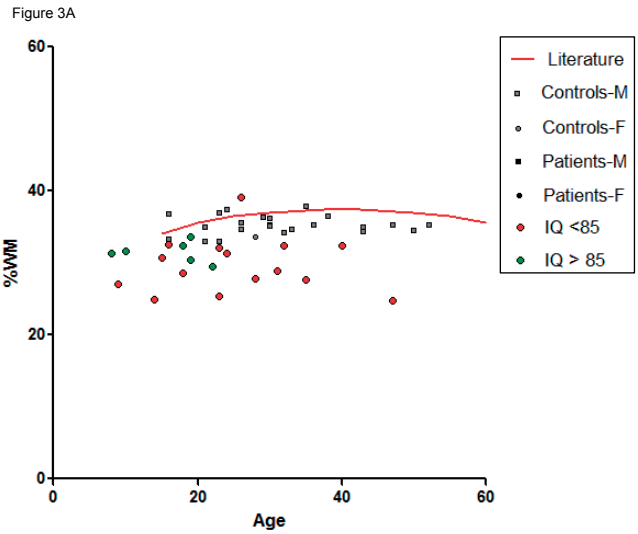


Figure 3A. Data from patients, controls and the literature based on healthy subjects (red line)[37].

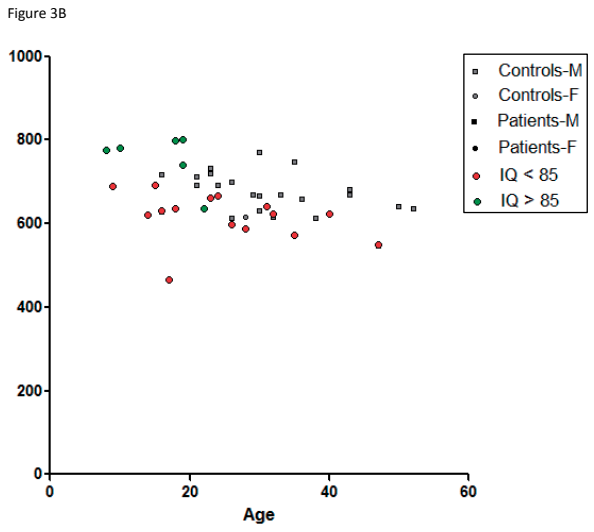


Figure 3B. Data from patients and controls.

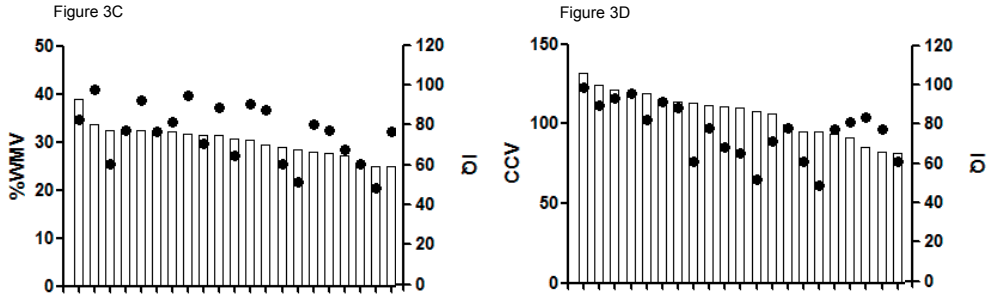


Figure 3C/3D. Every bar represents the %white matter volume (%WMV)/cerebral cortex volume (CCV) of one patient on the left Y-axis and the corresponding IQ on the right Y-axis.

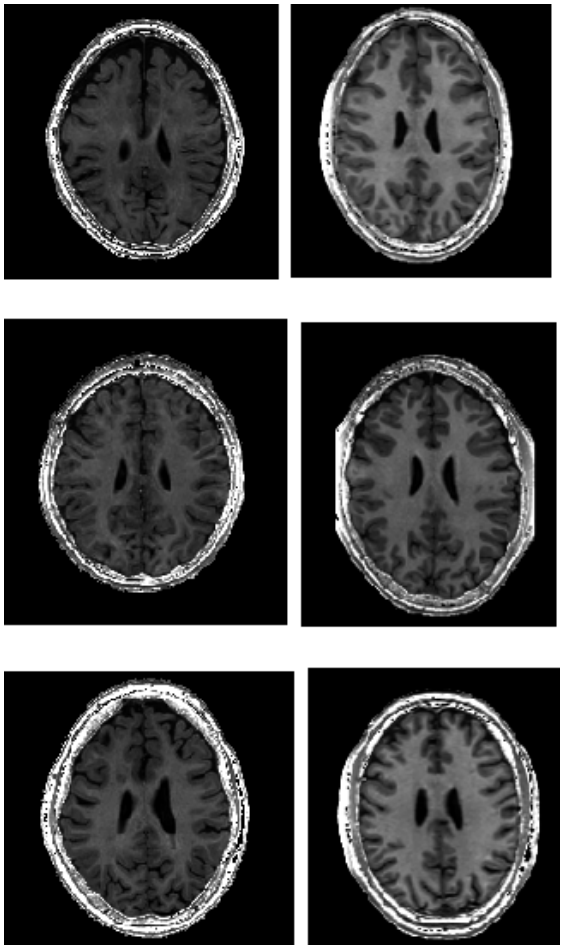


Figure 4. T1-weighted images of three age- and gender matched patients (left) and controls (right), illustrating the observed differences in white matter signal intensity between patients and controls.

DISCUSSION

This study investigated structural changes of the brain in CG patients with the use of MRI and its association with clinical outcome. The results of this study indicate that besides the white matter (WM), gray matter (GM) is also affected in CG patients. Patients with a poor neurological outcome, predominantly caused by the presence of a tremor, demonstrated more WM pathology than patients with a normal neurological outcome and the severity of the tremor was associated with several neuroimaging parameters. Patients with a poor intellectual outcome demonstrated more GM and WM pathology than patients with a normal intellectual outcome and the IQ was associated with several neuroimaging parameters.

In this study, we first investigated the differences in volumetric measurements, DTI measures and R1 values indicative of myelin content between patients and controls. CG patients demonstrated less WM volume when compared with controls. According to the literature, WM volume loss occurs in a non-linear, quadratic fashion from the fourth decade on[37]. In our cohort, patients demonstrated the same trend, but an overall lower absolute and fractional (%)WM volume than controls. The existence of WM pathology in CG patients as indicated by a lower WM volume was confirmed by the differences in DTI measures between patients and controls. The finding that FA was lower in CG patients in both the whole brain and in the corticospinal tract (CST) is in agreement with an earlier study which also demonstrated a lower density (NDI) and higher orientation dispersion (ODI) of neurites in several tracts[19]. As age influences DTI measures, with a decrease in FA and an increase in mean diffusivity (MD) and radial diffusivity (RD) with aging[38-40], and the patients in our cohort were younger and yet still demonstrated lower FA values than controls, our results may be an underestimation. Even though FA is a measure of microstructural integrity and highly sensitive to microstructural changes, it is less specific to the type of change[41]. To investigate this further, R1 values derived from qT1 scans were used to evaluate myelin content. Patients demonstrated lower R1 values of the WM, GM and CST when compared with controls, which could indicate that the data derived from DTI and qT1 reflect differences in myelination.

According to the literature, GM volume loss occurs in a linear fashion from late childhood on and males have a tendency to demonstrate a faster decrease in %GM volume than females[37]. As the patients in our cohort were younger and the included controls were predominantly males, more research is needed on the course of GM volume over time in CG patients. The comparable results between the younger patients and older, predominantly male controls suggests a lower GM volume in CG patients. In addition, the decline of GM volume in male patients appeared to be steeper when compared with male controls.

Second, volumetric measurements, DTI measures and R1 values were investigated between patients with a poor (movement disorders present) and normal

neurological outcome. Patients with a poor neurological outcome demonstrated a lower WM volume, an impaired WM microstructure in the whole brain and lower R1 values of WM when compared with patients with a normal neurological outcome. The association between neuroimaging parameters and the tremor rating scale indicated that patients with a more severe tremor had a lower WM volume in both the cerebrum and cerebellum, an impaired WM microstructure of the whole brain and corticospinal tract (CST) and less myelin content in both the GM and WM. The lower FA and higher radial diffusivity (RD) in patients with a more severe tremor could indicate less or damaged myelin.

Third, we investigated differences in volumetric measurements, DTI measures and R1 values between patients with a poor (IQ<85) and a normal (IQ≥85) intellectual outcome. Patients with a poor intellectual outcome demonstrated lower GM and WM volume in both the cerebrum and the cerebellum when compared with patients with a normal intellectual outcome, which was confirmed by the association between IQ as continuous measure and several neuroimaging parameters. Also, a lower IQ was associated with lower R1 values indicating less myelin content in both GM and WM.

According to the literature females generally have a higher %WM compared with males[37]. In our cohort, patients with an IQ<85 were predominantly females and the oldest patient with a poor neurological outcome was a female in her early thirties. As decline starts after the fourth decade, the pattern of WM volume over time in CG patients in general and more importantly in patients with long-term complications should be investigated further, as the results demonstrated in this study could be an underestimation. Importantly, the sub analysis performed in male patients where an age difference was no longer present, still demonstrated a lower GM and WM volume in both the cerebrum and cerebellum of patients with a poor intellectual outcome.

In this study, R1 values indicative of myelin content have been evaluated as previous studies suggested that myelin is affected in CG patients[2,5,14,15]. The formation of myelin depends on galactose containing glycolipids, such as galactocerebrosides[42,43]. The deficient GALT enzyme in CG patients causes a reduction in UDP-sugars which are essential for the formation of glycoproteins and glycolipids[8]. Therefore, the observed WM abnormalities on MRI in CG patients, such as the abnormal WM intensity signal, have been attributed to abnormal myelination. The finding that patients in our cohort demonstrated lower R1 values of the GM, WM and CST when compared with controls, supports the hypothesis that myelination is affected in CG patients. The association between R1 values of WM and GM and the severity of long-term complications, could indicate that myelin is more affected in patients with long-term complications.

Furthermore, we investigated the extent of WM hyperintensities. The volumes of WM hyperintensities were comparable between patients and controls, and between patients with and without long-term complications. Per decade, patients demonstrated a slightly higher volume of WM hyperintensities than controls, but the values observed

in patients were comparable to values reported in the literature for healthy subjects[44] and therefore WM hyperintensities were not considered clinically relevant and were not evaluated further in this study. These results demonstrate that visible WM hyperintensities on MRI are scarce and not as evident as one would expect in a disease primarily affecting WM as seen in Metachromatic leukodystrophy (MLD) and Multiple Sclerosis (MS). Interestingly, besides progressive demyelination and widespread visible WM abnormalities on MRI, both MLD and MS are WM disorders in which the GM is also affected[45,46].

The fact that the most common long-term complications in CG patients are cognitive impairment and movement disorders suggests that CG is a GM disease. In patients with movement disorders, WM was primarily affected and in patients with a poor intellectual outcome both GM and WM were affected. Also, widespread WM hyperintensities were absent in CG patients and WM hyperintensity volumes were comparable to controls, The results found in our study indicate both WM volume loss and GM volume loss in CG patients. In case of neuronal degeneration not only the cell body in the cortex degenerates, but also the accompanying supporting cells and WM tracts, leading to relatively more WM loss than GM loss, even though GM is primarily affected. The higher %GM found in patients could indicate a stronger decrease in WM compared to GM. Based on the clinical outcome of our patients and the result found in our study, we propose that CG is primarily a GM disease with secondary damage to the WM as a result of neuronal degeneration. If CG is a neurodegenerative disease, the CSF volume would be increased in patients when compared with controls. In our study, patients demonstrated a lower absolute CSF volume when compared with controls, which is not unexpected as the included patients were significantly younger than the included controls. Interestingly, the %CSF volume was comparable between patients and controls despite the age difference. As the severity of a tremor and dystonia and a lower IQ were associated with a higher CSF volume, the loss of brain volume as indicated by a higher CSF volume seems to contribute to the pathology of CG.

Strengths

The investigated neuroimaging parameters depend on the quality of the MR images and motion artifacts can influence the results. Therefore, analyses were repeated after the exclusion of patients with a lower quality. Even though this resulted in a smaller sample size, the overall trend remained. It is important to be aware that all excluded patients had a poor clinical outcome and some of the excluded patients were among the most severely affected in our cohort.

Besides the statistical analyses, the included graphical presentation of the data provides a more reliable insight on the distribution of WM and GM in patients with varying ages, in patients compared to controls and in patients with a poor and normal clinical outcome.

Limitations

Based on available literature, both age- and gender are important confounders when studying the brain on MRI. As our control group consisted of 23 males and only one female and the 21 included patients were significantly younger than the controls, this study should be considered as an explorative study and results should be interpreted with care. The effects of age and gender on neuroimaging parameters are a frequently studied subject, but available literature demonstrates contradicting results with regard to the magnitude of the effect. This could arise from the use of different methods as there is no gold standard and several analysis programs are available. Also, volume loss is heterogeneous and the investigated areas of the brain differ between studies and the distribution of age and gender differs between cohorts. The effects of age and gender as confounders on volumetric measurements and age on DTI measures were confirmed in our study and the sub-analyses of patients based on gender and long-term complications resulted in an even smaller sample size. A correction for age was not considered desirable in this small cohort. Correction for multiple testing was not performed due to the explorative nature of this study and due to the small sample size.

The effect of gender on DTI measures remains unclear and also depends on the investigated areas of the brain[38,40]. In our cohort DTI measures for the whole brain and the CST were not significantly different between males and females and therefore no sub-analyses were performed.

The abnormal white matter intensity signal as described by Nelson et al[15], was found in our cohort as well and was more evident in patients with severe long-term complications. The finding of an abnormal white matter signal intensity on the MRIs of some CG patients may have caused some bias in the segmentation of GM and WM because the analysis programs used for the volumetric measurements depend on the contrast between WM and GM to determine the volume.

In this study, FLAIR and T1 images were used to assess WM hyperintensities. For certain regions in the brain a T2-weighted sequence is more sensitive for WM hyperintensities and therefore could have been of added value.

Future perspectives

Considering this is a cross-sectional study, the question remains whether the structural changes of GM and WM observed in CG patients are a consequence of damage early on in life with an incomplete brain maturation/ dysmyelination or whether CG is a neurodegenerative disease with progressive damage throughout life with increasing volume loss and demyelination. To investigate this further, longitudinal studies are needed in a larger and age-and gender matched cohort, which could also provide insight in the course of GM and WM over time in CG patients.

As it remains uncertain if GM or WM is primarily affect in CG, more research is needed to explore our hypothesis. As the most severely affected patients demonstrated

an abnormal white matter signal intensity on MRI, quantitative analyses that are not based on contrast such as the quantitative T1 mapping may provide a more reliable reflection of GM and WM content and could be of use to gain knowledge on the pathophysiology of CG. Furthermore histopathological data could contribute highly to unravel the pathophysiology of CG.
In this study, IQ was used as derivative of intellectual outcome, but IQ only represents overall cognitive functioning and may not reflect underlying cognitive abilities[47]. Considering a link between WM pathology and cognitive functioning in CG patients has been demonstrated[19] and in our study IQ was not associated with DTI measures of the whole brain, but was associated with R1 values of GM and WM indicative of myelin content, the correlation between structural changes of the brain and neurocognitive outcome measures should be investigated further.

CONCLUSION

This study confirms WM pathology affecting the volume, microstructure and myelin content of the brain in CG patients. Besides WM pathology, this study demonstrates that GM was affected as well. Even though our cohort was small, considerable differences in neuroimaging parameters were demonstrated between patients and controls and between patients with a poor and normal clinical outcome. In patients with movement disorders, WM was primarily affected, whereas in patients with a poor intellectual outcome both GM and WM were affected. Based on the clinical outcome of patients, the absence of widespread WM hyperintensities and the finding that both GM and WM are affected, we propose that CG is primarily a GM disease with secondary damage to the WM as a result of neuronal degeneration. The finding of an association between structural changes of the brain and the severity of long-term complications supports the use of quantitative MRI techniques to further unravel the pathophysiology of CG. In addition, longitudinal research that reveals the course of GM and WM could clarify if CG is a neurodegenerative disease or whether damage already occurs at an early stage.

ACKNOWLEDGEMENTS

The authors would like to thank Sandra van den Berg and Raschel van Luijk for their assistance with the MRI settings and for conducting the MRI scans.

Funding
This study was supported by a grant of Stofwisselkracht.

REFERENCES

1. Bosch AM. Classical galactosaemia revisited. *Journal of inherited metabolic disease*. 2006;29(4):516-25.
2. Hughes J, Ryan S, Lambert D, Geoghegan O, Clark A, Rogers Y, et al. Outcomes of siblings with classical galactosemia. *Journal of pediatrics*. 2009;154(5):721-6.
3. Waisbren SE, Potter NL, Gordon CM, Green RC, Greenstein P, Gubbels CS, et al. The adult galactosemic phenotype. *Journal of inherited metabolic disease*. 2012;35(2):279-86.
4. Coss KP, Doran PP, Owoeye C, Codd MB, Hamid N, Mayne PD, et al. Classical Galactosaemia in Ireland: incidence, complications and outcomes of treatment. *Journal of inherited metabolic disease*. 2013;36(1):21-7.
5. Fridovich-Keil JL, Walter JH. Part 7: Carbohydrates, Chapter 72: Galactosemia.The Online Metabolic and Molecular Bases of Inherited Disease, OMMBID; Valle D.L., Antonarakis S, Ballabio A, Beaudet A.L., Mitchell G.A.(Eds.). McGraw Hill, New York.
6. Ng WG, Xu YK, Kaufman FR, Donnell GN. Deficit of uridine diphosphate galactose in galactosaemia. *Journal of inherited metabolic disease*. 1989;12(3):257-66.
7. Keevill NJ, Holton JB, Allen JT. The investigation of UDPGlucose and UDPGalactose concentration in red blood cells of patients with classical galactosaemia. *Clinica chimica acta*. 1993;221(1-2):135-42.
8. Lai K, Langley SD, Khwaja FW, Schmitt EW, Elsas LJ. GALT deficiency causes UDP-hexose deficit in human galactosemic cells. *Glycobiology.* 2003;13(4):285-94.
9. Lai K, Tang M, Yin X, Klapper H, Wierenga K, Elsas L. ARHI: A new target of galactose toxicity in Classic Galactosemia. *Bioscience Hypotheses.* 2008;1(5):263-71.
10. Coman DJ, Murray DW, Byrne JC, Rudd PM, Bagaglia PM, Doran PD, et al. Galactosemia, a single gene disorder with epigenetic consequences. *Pediatric research*. 2010;67(3):286-92.
11. Coss KP, Byrne JC, Coman DJ, Adamczyk B, Abrahams JL, Saldova R, et al. IgG N-glycans as potential biomarkers for determining galactose tolerance in Classical Galactosaemia. *Molecular genetics and metabolism*. 2012;105(2):212-20.
12. Maratha A, Stockmann H, Coss KP, Estela Rubio-Gozalbo M, Knerr I, Fitzgibbon M, et al. Classical galactosaemia: novel insights in IgG N-glycosylation and N-glycan biosynthesis. *European journal of human genetics*. 2016;24(7):976-84.
13. Colhoun HO, Rubio Gozalbo EM, Bosch AM, Knerr I, Dawson C, Brady J, et al. Fertility in classical galactosaemia, a study of N-glycan, hormonal and inflammatory gene interactions. *Orphanet journal of rare diseases*. 2018;13(1):164.
14. Koch TK, Schmidt KA, Wagstaff JE, Ng WG, Packman S. Neurologic complications in galactosemia. *Pediatric Neurology.* 1992;8(3):217-20.
15. Nelson MD, Jr., Wolff JA, Cross CA, Donnell GN, Kaufman FR. Galactosemia: evaluation with MR imaging. *Radiology.* 1992;184(1):255-61.
16. Otaduy MC, Leite CC, Lacerda MT, Costa MO, Arita F, Prado E, et al. Proton MR spectroscopy and imaging of a galactosemic patient before and after dietary treatment. *American Journal of Neuroradiology.* 2006;27(1):204-7.

17. Krabbi K, Uudelepp ML, Joost K, Zordania R, Ounap K. Long-term complications in Estonian galactosemia patients with a less strict lactose-free diet and metabolic control. *Molecular genetics and metabolism*. 2011;103(3):249-53.
18. Rubio-Agusti I, Carecchio M, Bhatia KP, Kojovic M, Parees I, Chandrashekar HS, et al. Movement disorders in adult patients with classical galactosemia. *Movement disorders.* 2013;28(6):804-10.
19. Timmers I, Zhang H, Bastiani M, Jansma BM, Roebroeck A, Rubio-Gozalbo ME. White matter microstructure pathology in classic galactosemia revealed by neurite orientation dispersion and density imaging. *Journal of inherited metabolic disease*. 2015;38(2):295-304.
20. Timmers I, van der Korput LD, Jansma BM, Rubio-Gozalbo ME. Grey matter density decreases as well as increases in patients with classic galactosemia: A voxel-based morphometry study. *Brain Research.* 2016;1648(Pt A):339-44.
21. Kaufman FR, McBride-Chang C, Manis FR, Wolff JA, Nelson MD. Cognitive functioning, neurologic status and brain imaging in classical galactosemia. *European journal of pediatrics*. 1995;154(7 Suppl 2):S2-5.
22. Huffnagel IC, van Ballegoij WJC, Vos J, Kemp S, Caan MWA, Engelen M. Longitudinal diffusion MRI as surrogate outcome measure for myelopathy in adrenoleukodystrophy. *Neurology.* 2019;93(23):e2133-e43.
23. Fahn S, Tolosa E, Marín C, tremor. Crsf. Parkinson's Disease and Movement Disorders. Jankovic J TE, editors., editor. Baltimore: : Williams & Wilkins; 1993.
24. Burke RE, Fahn S, Marsden CD, Bressman S, Moskowitz C, Friedman J. Validity and reliability of a rating scale for the primary torsion dystonias. *Neurology.* 1985;35(1):73-7.
25. Tournier JD, Smith R, Raffelt D, Tabbara R, Dhollander T, Pietsch M, et al. MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *Neuroimage.* 2019;202:116137.
26. Andersson JLR, Sotiropoulos SN. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage.* 2016;125:1063-78.
27. Andersson JLR, Graham MS, Zsoldos E, Sotiropoulos SN. Incorporating outlier detection and replacement into a non-parametric framework for movement and distortion correction of diffusion MR images. *Neuroimage.* 2016;141:556-72.
28. Bastiani M, Cottaar M, Fitzgibbon SP, Suri S, Alfaro-Almagro F, Sotiropoulos SN, et al. Automated quality control for within and between studies diffusion MRI data using a non-parametric framework for movement and distortion correction. *Neuroimage.* 2019;184:801-12.
29. Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, et al. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage.* 2006;31(4):1487-505.
30. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, et al. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage.* 2004;23 Suppl 1:S208-19.
31. Mori S, Wakana S, van Zijl P, Nagae-Poetscher LM. MRI Atlas of Human White Matter: Elsevier Science; 2005.

32. Wakana S, Caprihan A, Panzenboeck MM, Fallon JH, Perry M, Gollub RL, et al. Reproducibility of quantitative tractography methods applied to cerebral white matter. *Neuroimage.* 2007;36(3):630-44.
33. Hua K, Zhang J, Wakana S, Jiang H, Li X, Reich DS, et al. Tract probability maps in stereotaxic spaces: analyses of white matter anatomy and tract-specific quantification. *Neuroimage.* 2008;39(1):336-47.
34. Stuber C, Morawski M, Schafer A, Labadie C, Wahnert M, Leuze C, et al. Myelin and iron concentration in the human brain: a quantitative study of MRI contrast. *Neuroimage.* 2014;93 Pt 1:95-106.
35. Poot DH, Klein S. Detecting statistically significant differences in quantitative MRI experiments, applied to diffusion tensor imaging. *IEEE transactions on medical imaging*. 2015;34(5):1164-76.
36. Deoni SC. High-resolution T1 mapping of the brain at 3T with driven equilibrium single pulse observation of T1 with high-speed incorporation of RF field inhomogeneities (DESPOT1-HIFI). *Journal of Magnetic Resonance Imaging.* 2007;26(4):1106-11.
37. Ge Y, Grossman RI, Babb JS, Rabin ML, Mannon LJ, Kolson DL. Age-related total gray matter and white matter changes in normal adult brain. Part I: volumetric MR imaging analysis. *American Journal of Neuroradiology.* 2002;23(8):1327-33.
38. Hsu JL, Leemans A, Bai CH, Lee CH, Tsai YF, Chiu HC, et al. Gender differences and age-related white matter changes of the human brain: a diffusion tensor imaging study. *Neuroimage.* 2008;39(2):566-77.
39. Sala S, Agosta F, Pagani E, Copetti M, Comi G, Filippi M. Microstructural changes and atrophy in brain white matter tracts with aging. *Neurobiology of aging.* 2012;33(3):488-98 e2.
40. Rathee R, Rallabandi VP, Roy PK. Age-Related Differences in White Matter Integrity in Healthy Human Brain: Evidence from Structural MRI and Diffusion Tensor Imaging. *Magnetic resonance insights.* 2016;9:9-20.
41. Alexander AL, Lee JE, Lazar M, Field AS. Diffusion tensor imaging of the brain. *Neurotherapeutics.* 2007;4(3):316-29.
42. Petry K, Greinix HT, Nudelman E, Eisen H, Hakomori S, Levy HL, et al. Characterization of a novel biochemical abnormality in galactosemia: deficiency of glycolipids containing galactose or N-acetylgalactosamine and accumulation of precursors in brain and lymphocytes. *Biochemical medicine and metabolic biology*. 1991;46(1):93-104.
43. Lebea PJ, Pretorius PJ. The molecular relationship between deficient UDP-galactose uridyl transferase (GALT) and ceramide galactosyltransferase (CGT) enzyme function: a possible cause for poor long-term prognosis in classic galactosemia. *Medical hypotheses.* 2005;65(6):1051-7.
44. Pagani E, Agosta F, Rocca MA, Caputo D, Filippi M. Voxel-based analysis derived from fractional anisotropy images of white matter volume changes with aging. *Neuroimage.* 2008;41(3):657-67.
45. Calabrese M, Magliozzi R, Ciccarelli O, Geurts JJ, Reynolds R, Martin R. Exploring the origins of grey matter damage in multiple sclerosis. *Nature reviews neuroscience.* 2015;16(3):147-58.
46. Tillema JM, Derks MG, Pouwels PJ, de Graaf P, van Rappard DF, Barkhof F, et al. Volumetric MRI data correlate to disease severity in metachromatic leukodystrophy. *Annals of clinical and translational neurology.* 2015;2(9):932-40.

47. Welsink-Karssies MM, Oostrom KJ, Hermans ME, Hollak CEM, Janssen MCH, Langendonk JG, et al. Classical galactosemia: neuropsychological and psychosocial functioning beyond intellectual abilities. *Orphanet journal of rare diseases*. 2020;15(1):42.
48. Inano S, Takao H, Hayashi N, Abe O, Ohtomo K. Effects of age and gender on white matter integrity. AJNR Am J Neuroradiol. 2011;32(11):2103-9.

SUPPLEMENTARY MATERIAL

Table 1. Age- and gender matched males

	Male patients	Male controls	P-value
n	5	5	-
Age in years	19 (15-31)	21 (16-32)	0.598
Volumetric measurements			
- TIV in cm^3	1520 (1304 – 1644)	1501 (1404 – 1597)	0.602
- WMV	484 (400 – 533)	516 (485 – 525)	0.251
- GMV	740 (641 – 801)	691 (615 – 717)	0.175
- CSFV	305 (212 – 388)	285 (259 – 360)	0.917
- %WMV	30.7 (28.8 – 33.6)	34.2 (32.9 – 36.8)	**0.028**
- %GMV	48.7 (44.3 – 53.0)	44.7 (43.4 – 47.7)	0.076
- %CSFV	19.0 (16.3 – 26.8)	18.9 (18.4 – 22.5)	0.754
- CWMV in mL	34.2 (26.8 – 37.6)	30.5 (26.3 – 31.6)	0.347
- CCV in mL	115.4 (109.7 – 132.0)	118.9 (110.3 – 139.6)	0.602
Intellectual outcome	91 (61-98)	-	-
- IQ<85 (*n*)	2		
- IQ≥85 (*n*)	3		
Neurological outcome		-	-
- Movement disorder (*n*)	2		
- No movement disorder (*n*)	2		

Data reported in median (range). TIV: total intracranial volume, WMV: absolute white matter volume, GMV: absolute gray matter volume CSF: absolute cerebral spinal fluid volume, %WMV: relative white matter volume, %GMV: relative gray matter volume, %CSF: relative cerebral spinal fluid volume, CWMV: cerebellar white matter volume, CCV: cerebellar cortex volume, IQ: intellectual quotient.

Table 2.1. White matter hyper intensity volumes of patients

Age group (years)	N	M / F	Age	DWMH volume (ml)	PVWMH volume (ml)	Total WMH volume (ml)
0 -10	2	3 / -	9 (8-10)	0.00	0.00	0.00
10-20	7	5 / 2	18 (14-19)	0.00	0.15	0.15
21-30	6	- / 6	23.5 (22-28)	0.03	0.15	0.18
31-40	4	1 / 3	33.5 (31-40)	0.07	0.21	0.28
41-50	1	0 / -	47	0.25	0.06	0.31
51-60	-	-	-	-	-	-

Data reported in median (range). M: male, F: female, DWMH: deep white matter hyper intensity, PVWMH: periventricular white matter hyper intensity.

Table 2.2. White matter hyper intensity volumes of controls

Age group (years)	N	M / F	Age	DWMH volume (ml)	PVWMH volume (ml)	Total WMH volume (ml)
0 -10	-	-	-	-	-	-
10-20	2	2 / -	16	0.00	0.04	0.04
21-30	12	11/ 1	26 (21-30)	0.00	0.02	0.02
31-40	5	5 / -	35 (32-38)	0.00	0.03	0.03
41-50	4	4 / -	45 (43-50)	0.00	0.08	0.08
51-60	1	1 / -	52	0.45	0.28	0.73

Data reported in median (range). M: male, F: female, DWMH: deep white matter hyper intensity, PVWMH: periventricular white matter hyper intensity.

Figure 1A

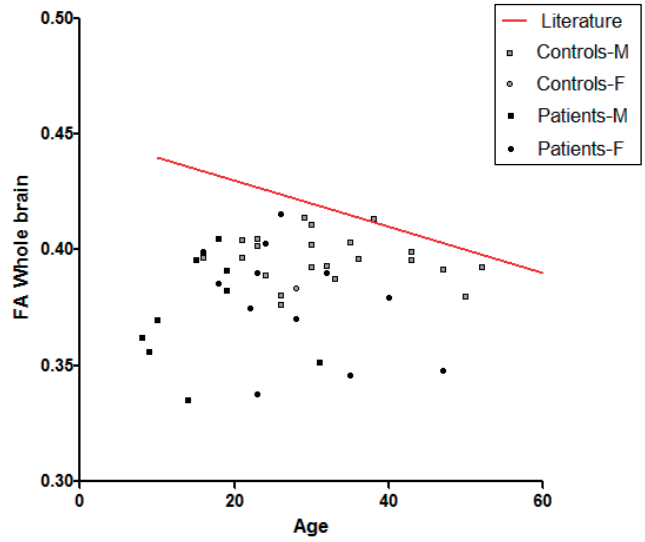


Figure 1A. Data from patients, controls and the literature based on healthy subjects (red line)[48].

Figure 1B

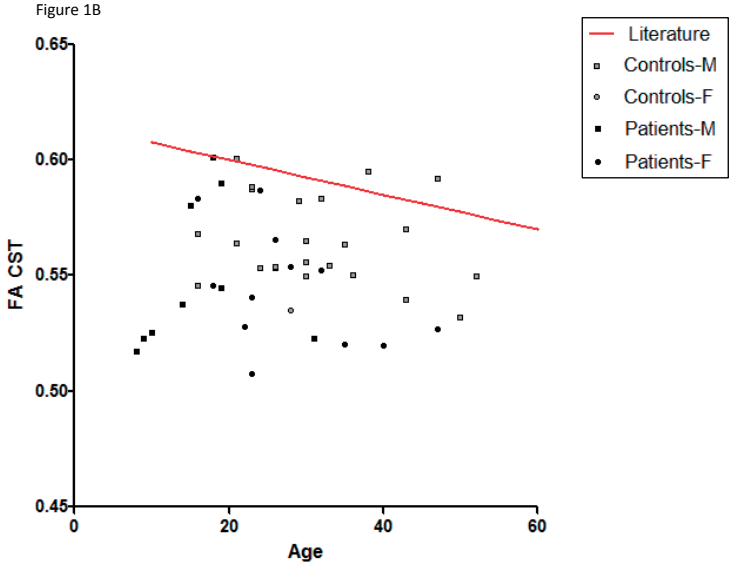


Figure 1B. Data from patients, controls and the literature based on healthy subjects (red line)[39].

Figure 1C

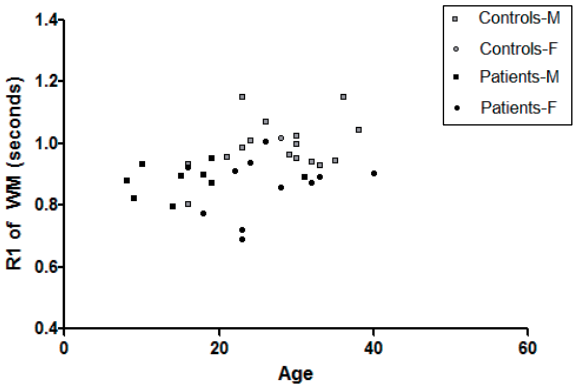


Figure 1C. Data from patients and controls.

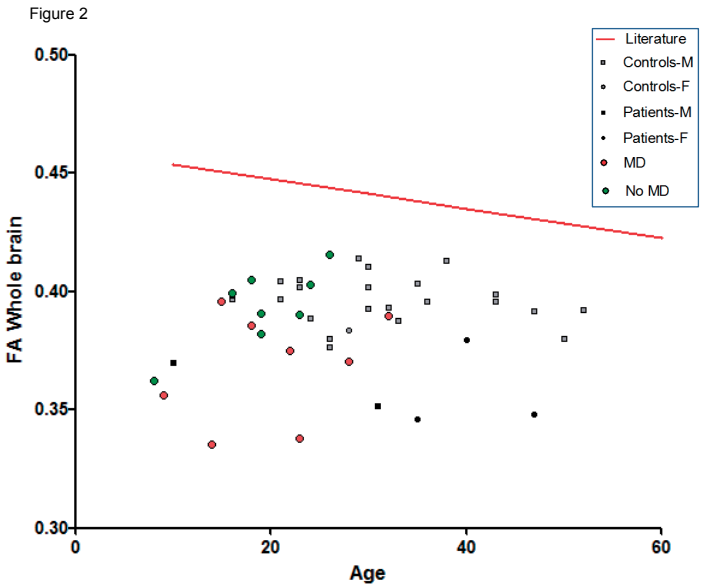


Figure 2. Data from patients, controls and the literature based on healthy subjects (red line)[48].

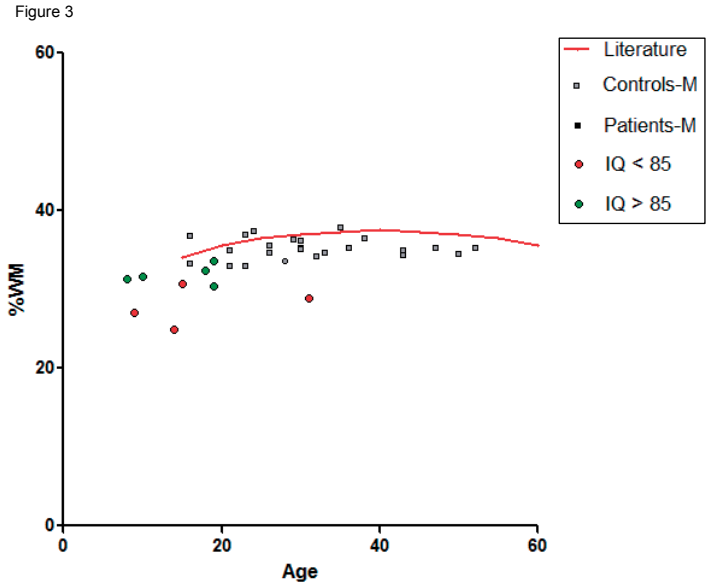


Figure 3. Data from patients, controls and the literature based on healthy subjects (red line)[37].

PART 3

PATIENT EDUCATION

Chapter 8

Reducing complexity: explaining inborn errors of metabolism and their treatment to children and adolescents

N.A. Zeltner, M.M. Welsink-Karssies, M.A. Landolt, D. Bosshard-Bullinger, F. Keller, A.M. Bosch, M. Groenendijk, S.C. Grünert, D. Karall, B. Rettenbacher, S. Scholl-Bürgi, M.R. Baumgartner, M. Huemer

Orphanet Journal of Rare Diseases. 2019 Nov;14(1):248.

ABSTRACT

Background

Inborn errors of metabolism (IEM) are a group of rare, heterogeneous and complex genetic conditions. Clinically, IEM often affect the central nervous system and other organs. Some carry the risk of progression and/or potentially life-threatening crises. Many patients have to adhere to lifelong dietary or drug treatment. The complexity of IEM makes it difficult for patients and caregivers to understand their pathophysiology, inheritance and therapy rationale. Especially patients reaching adolescence may have only limited knowledge of their condition since medical care has often entirely been handled by their parents. Knowledge about disease and treatment, however, constitute pillars of self-responsible disease management. Not many standardized patient education materials on IEM are available and their comprehensibility has not been systematically investigated.

Methods

We developed and tested patient education materials for school-aged children and adolescents with IEM. Informative texts and illustrations in paper form and as videos were developed by an international network of metabolic care professionals together with a graphic artist and experts for easy-to-read language. The materials were presented in standardized single or group training sessions to 111 individuals; first, to 74 healthy children and adolescents (recruited via public schools) and consecutively to 37 pediatric patients with IEM (phenylketonuria, galactosemia, urea cycle defects, lysosomal storage disorders) from six metabolic centers. Knowledge gain was assessed by pre- and post-testing.

Results

Knowledge gain was significant in healthy children and adolescents as well as in patients (p<0.001, r= -0.77 / -0.70). Effect sizes were large in both groups (r= -0.77 / -0.70). This result was independent from family language and teacher-rated concentration or cognitive capacity in healthy children.

Conclusion

The newly developed patient education materials are a powerful tool to improve disease- and treatment-related knowledge. They facilitate communication between the medical team and children and adolescents with IEM and their caregivers.

INTRODUCTION

Inborn errors of metabolism (IEM) are a group of rare, heterogeneous genetic conditions typically affecting enzyme functioning. Clinically, IEM often affect the central nervous system but also other organs. Some IEM carry the risk of "silent" progression while others cause phases of acute deterioration or even potentially life-threatening metabolic crises. Most patients carry a high burden of treatment, as they require lifelong dietary and/or drug treatment (e.g. ammonia scavengers, substrate inhibitors, enzyme replacement). Adherence to lifelong dietary or medical treatment is demanding and requires joint efforts of patients and families, healthcare providers, and institutions such as kindergarten or school.

Recently, questionnaires specifically assessing health-related quality of life (HrQoL) of patients with phenylketonuria[1] and intoxication type disorders[2] have been developed. Interviews and focus groups with patients and parents are integral parts of the construction process of HrQoL instruments and allow insight into patients' and parents' perspectives. In focus groups and interviews conducted by our research group[3], patients and parents communicated that the complexity and abstractness of IEM and their treatments made it almost impossible for them to understand the pathophysiology or treatment rationale - let alone explain this to relatives, friends or teachers. Consequently, they reported feelings of frustration, isolation and helplessness in social situations.

Comprehension of disease, treatment mechanisms and outcome determinants is an important pillar of successful, self-responsible disease management[4]. Education materials for children and adolescents with IEM are sparse and their quality and comprehensibility have never been investigated.

Processing fluency - defined as the subjective ease with which an individual is able to process new external information - constitutes an important determinant of patient motivation, self-management and adherence[5,6]. These factors are essential for optimal outcome and targets of patient education interventions[5,7]. Standardized materials for continuous comprehensive, attractive, industry-independent education about the disease and its treatment are an unmet medical need for pediatric patients with IEM[8,9].

It is well known that from age 10 years, dietary adherence in phenylketonuria (PKU) and acute intoxication-type disorders (e.g. urea cycle disorders, maple syrup urine disease) declines[10]. This effect may, at least in part, be attributed to both the cumbersome transition of health responsibility from parents to patients during adolescence and limited knowledge about disease and treatment[11,12].

This project aimed to develop and test specifically designed, standardized pictorial representations and videos combined with easy-to-understand texts for structured and comprehensible medical information on IEM for school-aged patients and their caregivers. We hypothesized that children and adolescents would be able to increase their knowledge about IEM after presentation of the materials.

METHODS

Development of materials

Materials were developed in cooperation with a graphic artist (BR) and specialists for easy to read language (Büro für Leichte Sprache, Lebenshilfe Bremen, Germany). Six metabolic physicians, two psychologists and one patient representative were involved in the development process. They defined the contents necessary to understand the basics of IEM and were involved in several feedback loops to evaluate and refine the paper materials and videos.

Materials

Initially, a total of 23 modules were developed to explain different types of IEM, their inheritance and treatment. Each module consists of 1-2 pages of pictorial presentations with short easy-to-read text elements. The testing phase in healthy children and adolescents showed that the material explaining that the body is built of cells, which contain genetic information, were too complex. Hence, the content was reformatted and presented in two steps, leading to a final set of 24 modules for the second testing phase in patients. Contents of the modules are listed in **Table 1**. **Figure 1** shows some examples (for more refer to[13]). Materials are available in English, Dutch, and German. Modules were constructed as to be freely combined by the metabolic team to explain different inheritance patterns, IEM of different types (e.g. storage, intoxication), as well as a variety of treatment options (e.g. scavengers, enzyme replacement).

Demographic data, abilities to concentrate and cognitive capacities

For all children and adolescents willing to participate, their parents completed a demographic questionnaire on basic information such as child age and family language. In the school setting, we asked teachers to rate concentration and cognitive capacities. Ability to concentrate was rated on a four-point Likert-scale. Cognitive capacities were assessed by using three items with a six-point Likert-scale indicating fluid intelligence of the THINK 1–4 (Baudson, Wollschläger & Preckel, 2016; Lehrpersonen-Einschätzungsskala zu den kognitiven Fähigkeiten; Test zur Erfassung der Intelligenz im Grundschulalter). This scale was chosen for its shortness and good internal consistency of Cronbachs $\alpha = 0.95$. Information about disease and school setting of the patients was part of the parent questionnaire.

Comprehensiveness testing

To evaluate the materials, they were assembled for each testing session to sets of pictures explaining normal body functions, pattern of inheritance, enzyme function and dysfunction as well as treatment principles (for example normal enzyme functioning, autosomal recessive inheritance, intoxication of cells, dietary treatment and metabolite

removal by drugs as a set for intoxication type disorders). Disease-related knowledge was assessed before and after the teaching session using 5 to 8 specifically developed multiple-choice items (number depending on content and age) and one open question. Test items of the knowledge test were developed by a metabolic physician, a psychologist, a psychology master student and a teacher. Piloting of the comprehensiveness test was performed in 41 children of four school classes, leading to a refinement of the test items and adjustment of complexity to avoid ceiling effects. Results of this pilot group are not included in the final analysis. After adaptation of the knowledge test, comprehensiveness testing was performed in healthy children and adolescents and later in patients. In both groups, materials were presented in a standardized teaching session given by trained persons with psychological and/or metabolic background.

Testing in healthy children and adolescents

Children and adolescents from 1st to 8th school grade (age 7 to 15 years) from nine classes in three public schools in the Zurich area (Switzerland) were invited to participate. Teachers distributed information material and consent forms to all children attending their class for discussion with their parents. Inclusion criteria were informed consent by parents and child, sufficient command of German and absence of IEM in the family. Hence, this group is referred to as the healthy group / healthy children and adolescents. Two members of the study group with psychological background performed the testing (30-45 min per group). Modules were assembled according to different case vignettes, which were presented to the groups. Consequently, all modules could be tested. Modules explaining inheritance were introduced only from the 5th grade upwards (>11 years) due to their complexity and abstractness. Before and directly after the teaching session, pre- and post-knowledge testing were performed.

Testing in patients

Convenience samples of patients with IEM from the metabolic clinics in Amsterdam (NL), Basel (CH), Bregenz (A), Freiburg (GE), Innsbruck (A) and Zurich (CH) were invited to participate. Patients were invited on the occasion of a scheduled visit at the outpatient clinic (all sites) and additionally during a patient group meeting (Amsterdam only). Testing was performed in individual patients or as a group testing with a maximum group size of eight patients. Pre- and post-knowledge was tested before and directly after a teaching session with the materials. Duration of the procedure varied between 25 to 50 minutes. Inclusion criteria were sufficient command of the German (study sites in Austria, Germany, and Switzerland) or Dutch (study site in the Netherlands) language and the ability to participate in the test situation.

Table 1. Education modules explaining IEM and their treatment

Main domain	Content		Available as print version*	Available as video sequence*
The body	The body is built of cells		x	
	Cells contain the "body construction plan"		x	
Patterns of inheritance	Autosomal recessive		x	
	Autosomal dominant		x	
	X-chromosomal recessive		x	
	X-chromosomal dominant		x	
The healthy body	Building blocks for the body from foods and other sources		x	x
	Enzymes at work in a healthy cell		x	x
Enzyme defects	…in intoxication-type diseases		x	x
	…in storage diseases		x	x
	…in diseases where important substances cannot be produced		x	x
Treatment	Medicine	…for intoxication-type diseases	x	x
		…for storage diseases	x	x
		…for diseases where important substances cannot be produced	x	x
	Diet	…for intoxication-type diseases	x	x
		…for storage diseases	x	
		…for diseases where important substances cannot be produced	x	x
	Enzyme helper	…for intoxication-type diseases	x	x
		…for storage diseases	x	x
		…for diseases where important substances cannot be produced	x	
	Enzyme replacement	…for intoxication-type diseases	x	x
		…for storage diseases	x	x
		…for diseases where important substances cannot be produced	x	
Emergency situations	Be careful! Recognize dangerous situations (vomiting, fever etc.)		x	x

* Video sequences were produced based on the print versions, which were tested in healthy children and adolescents as well as patients. All materials are available at[13].

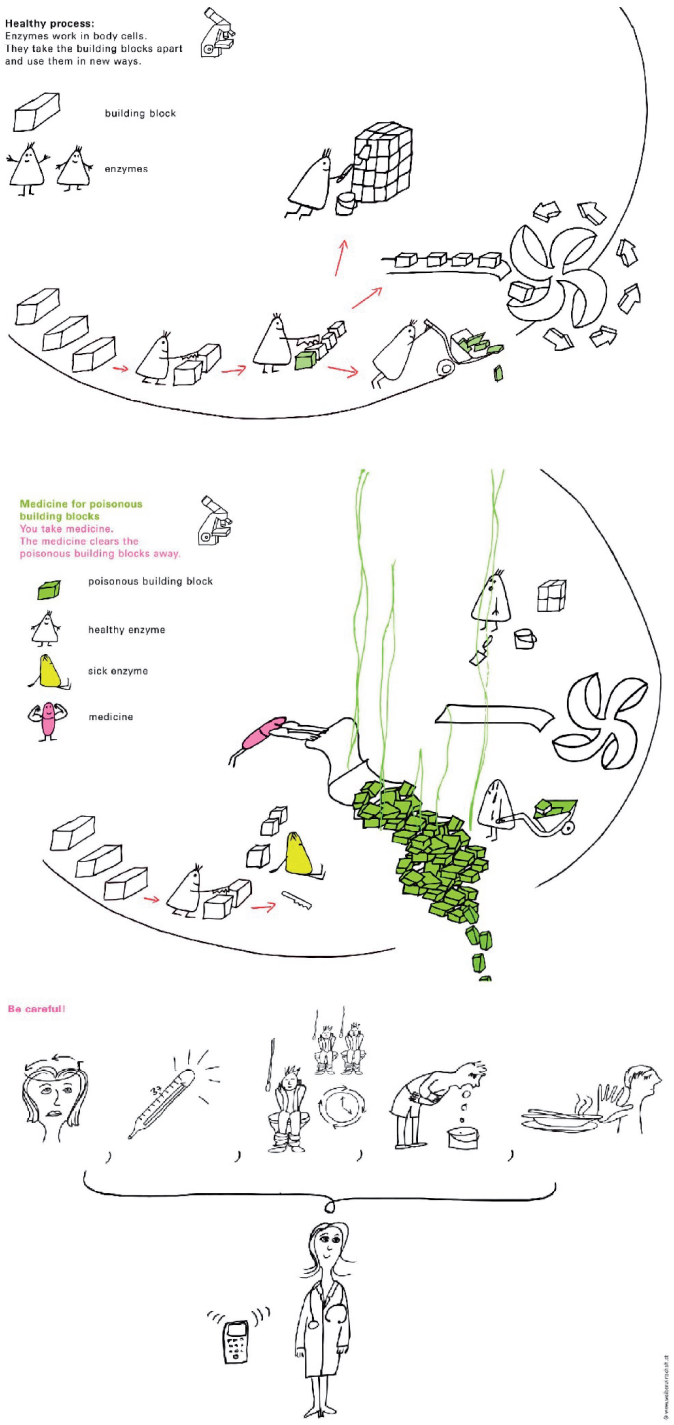


Figure 1. Example pictures (from top to bottom): (1) Enzymes at work in a healthy cell; (2) Medicine for intoxication-type disease; (3) Be careful! Recognise dangerous situations

Statistical analyses

Data from the healthy group and the patient group were analyzed independently, since materials were revised after the school-testing. Percent of correct answers were compared between pre- and post-testing to analyse knowledge gain in each of the two samples by Wilcoxon signed-rank tests. Answers to the open question counted double due to higher task complexity. Correlation coefficient r was calculated as an effect size measure for comparison of medians to define the magnitude of the difference between pre- and post-test results (0.10 small effect, 0.25 medium effect, 0.40 large effect)[14,15]. A significance level of $p<0.05$ was defined to indicate statistical significance. Participants with more than ≥50% missing answers in one test were excluded from this analysis.

Spearman's rank correlation was used to evaluate the relationship of knowledge gain and pre-test knowledge with potential influencing variables such as age, family language (same as language of the materials vs. other language), concentration and cognitive abilities (teacher's rating; only available for the healthy group), diagnosis of galactosemia (subgroup of patient group), school type (regular school without additional assistance vs. schooling with assistance; patient group only) and testing site (country). Galactosemia diagnosis was chosen as variable in the patient group due to the well-established high frequency of impaired cognitive abilities associated with it[16].

All statistical analyses were performed with the statistical software package SPSS, versions 24.0 (IBM Corp. IBM SPSS Statistics for Windows).

Subjective assessment of materials

Subjective assessment of materials ("how much do you like the materials?") was asked on a three-point smiley scale at the end of the post-test questionnaire. Participants were encouraged to contribute further comments on the materials by an open question.

RESULTS

Sample characteristics

One hundred and eleven children and adolescents were included in testing; 74 healthy children and adolescents (37 females; age range: 7.23–15.16; mean age: 11.56 ± 2.17) and 37 patients (19 females; age range: 6.96–19.18; mean age: 11.08 ± 3.17) with IEM (19 with phenylketonuria, 8 with galactosemia, 3 with propionic aciduria; 2 with hepatic glycogen storage diseases, 2 with Pompe disease, 1 each with mucopolysaccharidosis Type IVa, ornithin transcarbamylase deficiency and LCHAD). School setting (indicator for cognitive capacities) was regular school without additional assistance in 21 patients (57%) and schooling with assistance in 16 patients (43%). One patient was excluded from the analyses due to ≥50% of missing answers in the post-test, but included for the subjective assessment of the materials.

Testing of disease-related knowledge

Results of knowledge gain are depicted in **Figure 2**. Healthy children and adolescents showed a significantly higher percentage of correct answers in the post-test (median= 87.50) than in the pre-test (median= 37.50; z= -6.60, p<0.0005). This difference represents a large effect of r= -0.77. Patients also had a higher percentage of correct answers in the post-test (median= 66.67) compared with the pre-test (median= 37.50; z= -4.17, p<0.0005), which represents a large effect of r= -0.70.

Correlations of knowledge gain with different variables are reported in detail in **Table 2**. In the healthy group, knowledge gain was not significantly related to age, family language, ability to concentrate or cognitive capacities. In patients, no significant relation was found between knowledge gain and family language, school type, and galactosemia diagnosis. A significant negative correlation was found between knowledge gain and age and knowledge gain and pre-test knowledge, while pre-test knowledge and age showed a significant positive correlation.

Subjective assessment of materials on a three-point smiley scale, the majority of the healthy group (85%) and patients (84%) "liked" the materials, or rated them as "okay" (healthy group: 3%; patients: 14%). None of the healthy children and adolescents and only one patient disliked the materials (3%). Nine healthy children and adolescents (12%) did not answer this question.

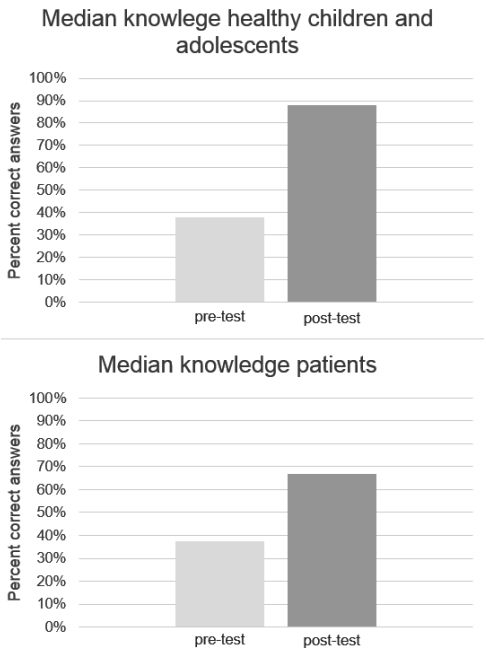


Figure 2. Knowledge gain in healthy children and adolescents and patients

Table 2. Bivariate correlations of knowledge gain and pre-test knowledge with potential influencing variables

Healthy children and adolescents

	1	2	3	4	5	6
1. Knowledge gain	-	-	-	-	-	-
2. Knowledge pre-test	**-0.77****	-	-	-	-	-
3. Age	-0.16	**0.46****	-	-	-	-
4. Family language	-0.19	0.11	0.18	-	-	-
5. Cognitive capacities	0.10	0.06	-0.05	-0.19	-	-
6. Ability to concentrate	0.15	0.00	-0.01	-0.09	**0.71****	-

Patients

	1	2	3	4	5	6	7
1. Knowledge gain	-	-	-	-	-	-	-
2. Knowledge pre-test	**-0.48****	-	-	-	-	-	-
3. Age	**-0.53****	**0.53****	-	-	-	-	-
4. Family language	0.04	-0.30	-0.12	-	-	-	-
5. Galactosemia	0.18	0.12	-0.32	0.19	-	-	-
6. School type	-0.13	-0.14	0.20	-0.03	-0.27	-	-
7. Testing site	-0.07	0.10	0.02	0.09	0.32	0.04	-

**Correlation is significant on the 0.01 level (2-tailed)

DISCUSSION

Use of unexplained medical jargon with families and patients is still common in the pediatric field[17] although it has been shown that information, which can be processed fluently, is considered more trustworthy, honest and safe and that the recipient is more inclined and likely to fulfil inferred tasks[5]. Processing fluency depends mainly on perceptive fluency, determined by visual display and linguistic fluency, which depends on wording and grammatical complexity. Our materials were designed to break down complicated concepts by pictorial representations and easy-to-read language texts to achieve good processing fluency for this complex information[5].

A main strength of our newly developed modular educational materials is that they have been thoroughly tested not only in patients from different countries, but also in naïve healthy children and adolescents. The materials can be used to educate IEM patients from school age and parents with any educational background and family language on different types of IEM (e.g. storage or intoxication type). They follow a modular system and thus allow for flexible combinations adaptable to individual cases, age groups and intellectual capacities, in a step-by-step approach.

Standardized teaching with our newly developed pictorial materials had a positive influence on disease-related knowledge in healthy children and in patients indicating good retrieval fluency of the materials. Knowledge gain was not related to age, family language and teacher's ratings of concentration abilities and cognitive capacities in the healthy group. Similarly, family language and indicators for cognitive capacities (galactosemia, school type) had no influence on knowledge gain in the patient group. However, knowledge gain was negatively related to age and pre-test-knowledge in patients. These results argue against our hypothesis that adolescents have little disease knowledge.

Experience during the testing sessions in healthy children showed that the rather abstract information e.g. about the body consisting of cells and genetics (DNA, genes, inheritance) were difficult to convey to younger children but worked well for adolescents. The contents how enzymes work and what happens when an enzyme is not working properly as well as treatment principles (e.g. diet, enzyme replacement) and information on how to recognize and act on emergencies were easier to grasp. Thus, materials about enzyme function, treatment options and how to act in emergency situations can be introduced in children from about 7 years of age while the more abstract topics (e.g. inheritance patterns) can be discussed in adolescence and/or with parents.

Although we controlled for potential effects of cognitive abilities on the understanding of the materials by different variables, a limitation of our study is that only information with regard to the attendance of special education was collected for the included patients. Therefore, we were not able to assess the effect of (severe) cognitive deficits on the test results. The short video sequences that were developed based on the cartoons may be

helpful for teaching both younger patients and patients with limited cognitive abilities, but this was not specifically tested. Furthermore, our teaching settings with groups of patients may not be representative for the usual clinical setting. We hypothesize, however, that the results should be replicable or may even be better in a one-to-one setting, which allows to fully focus on the needs of an individual / a family.

It is known that in patients with chronic diseases effective health education does not only improve knowledge, but also supports adherence to treatment and successful self-management, by influencing the patients' attitude and daily practice[17,18]. All these factors are essential for optimal outcome in patients with IEM[5,7]. In this study, so far, we only tested short-term knowledge gain. The effect of our materials on adherence, daily practice and biochemical outcome markers will be investigated in a next step. To allow wider use of the materials they will be accessible for professionals, patients and their families on an attractive patient-oriented homepage, which is another focus of this ongoing project.

Medical jargon and incomprehensive disease information constitute a barrier to informed and shared decision-making[17]. Over 80% of the study participants indicated that they liked the materials and appreciated them as an indicator of a respectful approach of care. Facilitated communication between patient and metabolic care team may not only enhance patients' safety, outcome, attitude towards the care team and quality of life but does significantly reduce the workload of metabolic professionals due to increased independence of patients and families[5]. Involvement of parents in teaching sessions and provision of the materials enables parents to explain the contents repeatedly to their child at home and to involve siblings, other family members or external caregivers to gain social support.

CONCLUSION

The newly developed materials are a powerful tool to enhance disease-related knowledge and processing fluency in children and adolescents with IEM from school age, preparing the ground for participative, safe and easy communication between patients and the metabolic team.

ACKNOWLEDGEMENTS

We would like to express our special thanks to all children, adolescents, parents, and teachers who participated in the study. We gratefully acknowledge the valuable contribution of Ariane Schweizer and Noemi Signer to the design of the knowledge test and to data management, respectively. We thank the language specialists from

Lebenshilfe Bremen who translated our texts into easy-to-read language.

Funding

We thank the Philhuman Stiftung, the Clinical Research Priority Program for Rare Diseases (radiz) at the University of Zurich and Nutricia – Advanced Medical Nutrition, Friedrichsdorf, Germany for financial support of this study.

REFERENCES

1. Regnault A, Burlina A, Cunningham A, Bettiol E, Moreau-Stucker F, Benmedjahed K, et al. Development and psychometric validation of measures to assess the impact of phenylketonuria and its dietary treatment on patients' and parents' quality of life: the phenylketonuria - quality of life (PKU-QOL) questionnaires. *Orphanet journal of rare diseases*. 2015; 10:59.
2. Zeltner NA, Baumgartner MR, Ensenauer R, Karall D, Kölker S, Mühlhausen C, et al. Development and psychometric evaluation of the MetaQOL 1.0 – a quality of life questionnaire for paediatric patients with intoxication-type inborn errors of metabolism. *JIMD reports*. 2017;37:27–35.
3. Zeltner NA, Landolt MA, Baumgartner MR, Lageder S, Quitmann J, Sommer R, et al. Living with Intoxication-Type Inborn Errors of Metabolism: A Qualitative Analysis of Interviews with Paediatric Patients and Their Parents. *JIMD reports*. 2016;31:1–9.
4. Boyd M, Lasserson TJ, McKean MC, Gibson PG, Ducharme FM, Haby M. Interventions for educating children who are at risk of asthma-related emergency department attendance. *Cochrane database systematic review*. 2009;CD001290.
5. Okuhara T, Ishikawa H, Okada M, Kato M, Kiuchi T. Designing persuasive health materials using processing fluency: a literature review. *BMC research notes*. 2017;10:198.
6. Rotliman AJ, Schwarz N. Constructing Perceptions of Vulnerability: Personal Relevance and the Use of Experiential Information in Health Judgments. Personality and Social Psychology Bulletin. Sage Publications Sage CA: Thousand Oaks, CA; 1998;24:1053–64.
7. ten Hoedt AE, Maurice-Stam H, Boelen CCA, Rubio-Gozalbo ME, van Spronsen FJ, Wijburg FA, et al. Parenting a child with phenylketonuria or galactosemia: implications for health-related quality of life. *Journal of inherited metabolic disease*. 2011;34:391–8.
8. Bregnballe V, Boisen K, Schiøtz P, Pressler T, Lomborg K. Flying the nest: a challenge for young adults with cystic fibrosis and their parents. *Patient preference and adherence*. 2017;Volume 11:229–36.
9. Steiß J-O, Lindemann H, Brosig B, Zimmer KP. Important aspects in pediatric care of children and adolescents with chronic disease using the example of bronchial asthma. *Deutsche Medizinische Wochenschrift*. 2013;138:2613–2618.
10. MacDonald A, van Rijn M, Feillet F, Lund AM, Bernstein L, Bosch AM, et al. Adherence issues in inherited metabolic disorders treated by low natural protein diets. *Annals of nutrition and metabolism*. 2012;61:289–95.
11. Crone MR, van Spronsen FJ, Oudshoorn K, Bekhof J, van Rijn G, Verkerk PH. Behavioural factors related to metabolic control in patients with phenylketonuria. *Journal of inherited metabolic disease*. 2005;28:627–637.
12. Mundy H, Lilburn M, Cousins A, Lee P. Dietary control of phenylketonuria. *The Lancet*. 2002;360:2076.
13. Education materials explaining inborn errors of metabolism. www.kispi.uzh.ch/fzk/stoffwechselpsychologie.
14. Field AP. Is the meta-analysis of correlation coefficients accurate when population correlations vary? *Psychological methods*. 2005;10:444–67.

15. Cohen J. Statistical Power Analysis. *Current Directions in Psychological Science*. SAGE Publications Sage CA: Los Angeles, CA; 1992;1:98–101.
16. Welling L, Waisbren SE, Antshel KM, Colhoun H-O, Gautschi M, Grünewald S, et al. Systematic review and meta-analysis of intelligence quotient in early-treated individuals with classical Galactosemia. *JIMD reports*. 2017. p. 115–123.
17. Connan V, Marcon MA, Mahmud FH, Assor E, Martincevic I, Bandsma RH, et al. Online education for gluten-free diet teaching: Development and usability testing of an e-learning module for children with concurrent celiac disease and type 1 diabetes. *Pediatric diabetes*. 2019;20:293–303.
18. Chawla SS, Kaur S, Bharti A, Garg R, Kaur M, Soin D, et al. Impact of health education on knowledge, attitude, practices and glycemic control in type 2 diabetes mellitus. *Journal of Family Medicine and Primary Care*. 2019;8:261.

TO CONCLUDE

Chapter 9

General discussion and future perspectives

GENERAL DISCUSSION

Background

Classical Galactosemia (CG) is an inborn error of galactose metabolism, which if left untreated, leads to a severe and often fatal disease affecting multiple organs in the neonatal period. An early diagnosis as a result of newborn- and family screening and hereby an early initiation of dietary galactose-restriction effectively treats neonatal illness. Unfortunately, an early treatment does not prevent long-term complications that primarily affect the brain and in most female patients the ovaries. The long-term outcome of patients is remarkably variable and prognostic biomarkers are currently lacking.

The pathophysiology of CG is not fully understood, but a severe deficiency of the galactose-1-phosphate uridylyltransferase (GALT) enzyme in CG patients leads to the accumulation of toxic metabolites (galactose-1-phosphate (Gal-1-P) and galactitol) prior to the GALT enzyme and the reduction of important metabolites downstream of the GALT enzyme, such as UDP sugars which are essential for the galactosylation of proteins and lipids. Both mechanisms are presumed to contribute to the long-term complications observed in CG patients[1,2].

The aim of this thesis is to elucidate the spectrum of clinical outcome in CG patients and to investigate potential (prognostic) markers for clinical outcome.

Patient groups and outcomes

CG is diagnosed if the GALT enzyme activity in erythrocytes is below 15% of the reference mean and/or if two pathogenic variations in the *GALT* gene are found. Until now, 363 variations in the *GALT* gene have been discovered, but a clear genotype-phenotype correlation in CG is lacking. In **Chapter 2**, the clinical and biochemical outcomes of the patients in our cohort were investigated. The prevalence of the most common long-term complications in CG patients (*i.e.* cognitive impairment, movement disorders and in females primary ovarian insufficiency, POI) were investigated. The investigated biochemical outcomes included erythrocyte Gal-1-P and urinary galactitol, two metabolites that are considered harmful and are presumed to contribute to the long-term complications observed in CG. In our cohort, we distinguished three groups of patients with differences in clinical and biochemical outcomes. The largest group of patients (n=47) are the patients with two classical pathogenic variations in the *GALT* gene resulting in low to undetectable GALT enzyme activity in erythrocytes (<3%) (hereafter named classical patients). Many patients suffer from long-term complications, but the clinical outcome within this group varies considerably, even in patients (and siblings) with identical disease-causing variations. We found a below average IQ (<85) in 68% of the patients (mean IQ 77, range 45-103), a movement disorder in 59% (a tremor and/or dystonia) and POI in 80% of the females aged 12 years and older. In classical

patients, levels of Gal-1-P and galactitol remained elevated despite dietary galactose restriction. A second group of patients (*n*=7), only detected since the introduction of CG in the newborn screening (NBS) program (in the Netherlands since 2007), demonstrates previously unreported genotypes and residual GALT enzyme activity up to 10% in erythrocytes[3]. During dietary galactose restriction, these patients (hereafter named variant patients) demonstrated undetectable Gal-1-P levels and normalized galactitol levels. The third group consists of two patients with a specific *GALT* variation, the homozygous p.Ser135Leu variation. These patients have a GALT deficiency in erythrocytes but residual GALT enzyme activity in other tissues[4]. Both patients (hereafter named homozygous p.Ser135Leu patients) were born before the introduction of CG in the Dutch NBS and were diagnosed late after they presented with CG related illness at age 7 months and 10 years. The presence of residual GALT enzyme activity may have protected them from severe neonatal illness, but both patients have a poor intellectual outcome with an IQ well below 85. As both patients have a normal neurological and endocrinological outcome, the question remains if the intellectual outcome of these patients would have been better if they were detected and treated early. In these patients, Gal-1-P and galactitol are lower when compared with classical patients but not fully normal as in the variant patients.

All variant patients in our cohort were detected by NBS and treated early with a galactose-restricted diet. As they did not demonstrate symptoms at diagnosis, the question remains if these individuals would have developed symptoms and thus would have been detected without NBS. As these patients were all diagnosed since 2007, they are still young (the oldest patients are now 9 years old) but currently all of them demonstrate a normal intellectual and neurological outcome. In contrast, many of the classical patients demonstrate motor- and speech delay already at an early age. Also, with the implementation of the International guideline advising neurological and neuropsychological evaluation from the age of 2-3 years[5], cognitive impairment and movement disorders are diagnosed early in classical patients. Therefore, it could very well be that variant patients treated early with dietary galactose restriction will not develop long-term complications.

Neuropsychological outcome

In **Chapters 3** and **4**, we investigated the neuropsychological outcome of CG patients. First, we performed a systematic review of the available literature in **Chapter 3**. Most studies that investigated cognitive functioning of patients primarily used intelligence tests without investigating the underlying cognitive domains. The studies that did investigated cognitive domains demonstrated different results, addressed only one cognitive domain, used a single test per domain and/or included small cohorts. Moreover, the quality of most studies was moderate to low. Therefore, we performed a comprehensive neuropsychological assessment in our patient cohort, which included multiple cognitive

functioning tests on several domains with the addition of questionnaires evaluating behavior and social functioning. The results described in **Chapter 4** demonstrate a large variability in cognitive, behavioral and social functioning in CG patients, which underlines the need for a standardized neuropsychological assessment in all patients. Overall, patients demonstrated a below average intelligence (mean IQ 77, range 45-103) and impaired cognitive functioning without a clear cognitive profile. The finding that cognitive outcomes and the educational level of patients were higher than expected based on their IQ may indicate that the IQ does not sufficiently reflect the abilities of patients. In our cohort, certain aspects of social functioning such as social responsiveness were investigated. The finding that social responsiveness of patients was comparable to the general population is important, as it has been previously suggested that CG patients demonstrate autistic traits. Considering internalizing problems such as social problems, withdrawn behavior and anxiety were frequently reported by (parents of) patients, more research into social functioning in CG patients is needed. The results from the neuropsychological evaluation will help to support patients based on their individual strengths and weaknesses, but it should be acknowledged that besides neuropsychological functioning, other factors such as adaptive functioning, coping and the support system are important and should be included as well.

Prognostic markers

In our search for potential prognostic markers for clinical outcome, we hypothesized that differences in outcome are caused by small differences in residual GALT enzyme activity leading to differences in residual galactose oxidation capacity. Individual differences in galactose oxidation capacity will affect the biochemical outcome (*i.e.* the extent of galactose intoxication as indicated by Gal-1-P levels and galactosylation abnormalities as indicated by *N*-glycan profiles) and may contribute to the observed differences in clinical outcome of CG patients. Therefore, we investigated Gal-1-P, *N*-glycans and galactose oxidation capacity, and their association with clinical outcome. The clinical outcome comprises the intellectual outcome (IQ</≥85), neurological outcome (presence or absence of movement disorders and in female patients aged 12 years and older, the endocrinological outcome (presence or absence of POI).

Gal-1-P

There is an ongoing debate on the predictive value of erythrocyte Gal-1-P and some studies have demonstrated an association with long-term outcome[6-8]. Therefore, in **Chapter 2**, we investigated the most recent Gal-1-P level in our cohort of patients and found that Gal-1-P was not able to differentiate between patients with a poor and normal clinical outcome. Interestingly, Gal-1-P was able to differentiate between classical patients and variant patients as variant patients demonstrated significantly lower Gal-1-P levels compared with classical patients. The differences in Gal-1-P levels

within the group of classical patients were small. As the most recent Gal-1-P level may not reliably reflect overall galactose intoxication, lifetime Gal-1-P was investigated in our cohort of classical patients as well. After the initiation of a galactose-restricted diet, Gal-1-P rapidly declines and is considered to reach a steady state within a year. Therefore, lifetime Gal-1-P was defined as the average of all Gal-1-P levels after the age of 12 months. In contrast to the findings of Yuzyuk et al[8], in our cohort lifetime Gal-1-P was not significantly different between patients with a poor and normal clinical outcome. Moreover, lifetime Gal-1-P was negatively correlated with age, indicating a more gradual but further decline over time. Therefore, it remains unclear if Gal-1-P reaches a steady state in patients, which is crucial for its use as prognostic biomarker. In addition, the use of a marker that is influenced by the extent to which a patient adheres to the diet may not be desirable.

N-glycan profiles

In our cohort, *N*-glycan abnormalities were investigated by analyzing the *N*-glycans on Immunoglobulin G (IgG). The *N*-glycan profiles consist of IgG *N*-glycan peaks, main glycans and *N*-glycan features[9-12]. Even though significant differences were found between the *N*-glycan profiles of patients and controls, no significant differences were found between patients with a poor and normal clinical outcome.

The finding that both Gal-1-P and *N*-glycan profiles were not associated with clinical outcome could be due to the fact that these potential markers were measured in erythrocytes and serum respectively and may not reflect Gal-1-P levels and *N*-glycans in affected tissues such as the brain and ovaries.

Galactose oxidation capacity

In **Chapter 5** and **6**, we investigated residual galactose oxidation capacity and the association with clinical outcome. In **Chapter 5**, we performed metabolic profiling in cultured skin fibroblasts by measuring labeled Gal-1-P and UDP-galactose. As Gal-1-P decreases and UDP-galactose increases in patients with higher galactose oxidation capacity, the Galactose Index (GI) (*i.e.* the ratio of [U^{13}C]Gal-1-P / [$^{13}C_6$]UDP-galactose) was calculated and used as indicator of galactose oxidation capacity.

The variant patients demonstrated a significantly lower GI (1.6 – 7.2) and thus a higher galactose oxidation capacity when compared with classical patients (9.0 – 22.4). Within the group of classical patients, the GI differed between patients but was not able to differentiate between patients with a poor and normal clinical outcome. The GI of the homozygous p.Ser135Leu patients was in the low range (2.5-5.4) and lower than some of the variant patients, indicating a higher galactose oxidation capacity in fibroblasts.

In **Chapter 6**, we measured whole body oxidation capacity by measuring labeled galactose in breath samples after the administration of 7 mg/kg of labeled galactose. The variant patients with lower Gal-1-P levels, higher residual GALT activity in erythrocytes

and higher galactose oxidation capacity in fibroblasts compared with classical patients, demonstrated significantly higher whole body galactose oxidation capacity (≥12) than classical patients (≤2). The classical patients demonstrated whole body galactose oxidation capacities in the low range and as the differences between patients were small, they were not able to differentiate between classical patients with a poor and normal clinical outcome. Both homozygous p.Ser135Leu patients with residual GALT activity in different tissues demonstrated a whole body galactose oxidation capacity in the control range (≥9).

The galactose oxidation capacity of the variant patients and the finding that all variant patients until now demonstrate a normal clinical outcome during treatment with galactose restriction is reassuring and raises the question if these patients are true patients in need of dietary treatment. Most of the younger patients in our cohort demonstrated a higher galactose oxidation capacity than the older patients with identical genotypes. As age-dependent GALT activity has been demonstrated in different tissues[13], more research is needed to clarify if the results of the young variant patients are influenced by their age or if the results truly reflect the galactose oxidation capacity of these patient. Because the whole body galactose oxidation capacity of the variant patients even exceeded the results of the included controls, it is important to consider if other factors that vary with age such as body composition, could have influenced the results of the galactose breath test. Interestingly, within the group of classical patients with a highly variable clinical outcome, whole body galactose oxidation was consistently in the low range as expected (≤2), whereas the galactose oxidation capacity in cultured skin fibroblasts varied considerably (GI ranging from 9.0 – 22.4). This might be due to tissue specific GALT activity, which has been observed in an animal model[13] and which may also explain why some organs are more affected than others. The finding that the GI is not associated with clinical outcome could indicate that cultured skin fibroblast are a poor reflection of residual galactose metabolism in other cells. The fact that we did not find a correlation between clinical and biochemical outcomes and residual galactose oxidation capacity in the group of classical patients could be due to our small cohort or may suggest that the differences in clinical outcome in this group may be attributed to other factors than residual galactose metabolism. Also, there may be a certain threshold for residual galactose oxidation capacity that is sufficient to provide a normal clinical outcome. The homozygous p.Ser135Leu patients demonstrated residual galactose oxidation capacity in cultured skin fibroblasts (GI ranging from 2.5 – 5.4), which is in agreement with the residual GALT activity found in other tissues than erythrocytes and could explain the whole body galactose oxidation capacity of these patients, which was in the normal range. The results of the galactose oxidation tests confirm the residual galactose metabolism in these patients, which could be sufficient to prevent long-term complications if treated early.

The finding that the sibling pairs of classical patients in our cohort demonstrated

comparable Gal-1-P levels, *N*-glycan profiles and residual oxidation capacity even though some of these siblings can be found on opposite ends of the clinical outcome spectrum, implies that other mechanisms than the above mentioned (bio)markers influence the clinical outcome of these patients.

Elucidating a new patient group

The findings of this thesis confirm that the variant patients are clinically and biochemically different from the classical patients and more importantly, they demonstrated residual galactose oxidation capacity in cultured skin fibroblasts and a whole body galactose oxidation capacity in the normal range. In 2019, data were published confirming that individuals with the Duarte Galactosemia genotype (with erythrocyte GALT activity between 14 and 25% and galactose oxidation capacity in the normal range) even without dietary treatment demonstrated similar developmental outcomes as their unaffected siblings which confirmed that these subjects are not in need of treatment and should not be considered as patients[14]. Currently, all patients with erythrocyte GALT enzyme activities below 15% are considered as patients and treated as such[5]. The variant patients in our cohort were detected with NBS since 2007 and their long-term follow-up has to be awaited, but the findings in this thesis indicate that their capacity to metabolize galactose may be sufficient to prevent long-term complications with current dietary treatment. It is not clear yet if some variant patients might benefit from a less strict diet, as galactose over restriction could harm patients as well[15]. Under strict biochemical control, some of our variant patients received dietary relaxation with an allowance up to 2400 mg galactose/day (100 mL of dairy products). Importantly, this did not increase erythrocyte Gal-1-P or urinary galactitol levels. Ultimately, it might become clear that some of these patients should not be considered patients at all but only have a biochemical variation.

Because the galactose oxidation capacity of these patients in fibroblasts varied, residual GALT activity in erythrocytes varied between 3 – 10% and the results of the galactose breath test could have been influenced by other factors, individual galactose tolerance should be investigated and closely monitored if dietary relaxation is considered. Moreover, the clinical outcome of these patients over time (and in comparison with their unaffected siblings) should be monitored. For now, these patients are advised to adhere to the diet as recommended by the guideline.

The use of MRI to unravel the pathology of CG

Since motor- and speech developmental delay, movement disorders and cognitive impairment are the most common long-term complications in CG, we studied the brain of CG patients with the use of conventional MRI, Diffusion Tensor Imaging (DTI) and quantitative MRI techniques. The aim was to investigate the association between abnormalities of the brain on MRI and clinical outcome. The qualitative

analyses performed in **Chapter 2** demonstrated MRI abnormalities in almost the entire cohort. Most patients demonstrated atrophy of the cerebrum and cerebellum, which was associated with movement disorders and a lower IQ. As pathological processes may be present in normal appearing white- and gray matter, in **Chapter 7**, we used a more quantitative approach to investigate structural changes of the brain on MRI and the association with clinical outcome. First, we investigated gray- and white matter volume, white matter integrity and myelin content in patients and controls. The results indicated a lower gray- and white matter volume, a lower white matter integrity and less myelin content in patients when compared with controls. Hereafter, we investigated if gray- and white matter volume, white matter integrity and myelin content differed between patients with a poor and normal intellectual and neurological outcome. Not only did patients with a poor clinical outcome demonstrated more structural changes of the brain, these structural changes were also associated with the severity of long-term complications such as the IQ and movement disorders. Based on the results, we propose that CG is primarily a disease affecting gray matter with secondary changes to the white matter as a result of neuronal degeneration. As age and gender are an important confounder for neuroimaging parameters, the results found in our study should be investigated further in a larger and age- and gender matched cohort, in which the aging brain of healthy subjects is compared with the aging brain of CG patients.
The MRIs of the severely affected CG patients in our cohort with IQs in the low range and/or severe movement disorders, demonstrated a remarkably abnormal white matter signal intensity as was previously described by Nelson et al[16]. The use of quantitative MRI techniques could be of use to further unravel the pathophysiology of CG.

In the final part of this thesis, we investigated the effectiveness of newly developed patient education materials. In **Chapter 8**, we found that these materials improved knowledge on disease and treatment in patients with CG and other inborn errors of metabolism (IEMs). As cognitive impairment is a frequently reported long-term complications in CG and other IEMs, it is important to educate patients on their own level and the developed materials may contribute to this.

FUTURE PERSPECTIVES

In this thesis, we explored the clinical outcome spectrum of CG and investigated potential (prognostic) markers for clinical outcome. The main findings of this thesis reflect the challenges in CG and underline the complexity of IEMs. As in many IEMs, there is an ongoing search for valid biomarkers in CG that will enable prognostication, individualized treatment and the development of new therapeutic options. Prognostication is urgently needed in all patients to provide clarity for parents and patients, but in variant patients with a different clinical and biochemical outcome, the finding of valid prognostic biomarkers may alter their treatment and follow-up.

Due to the rarity of the disorder, (international) collaboration is essential for future research. Potential biomarkers may demonstrate only small differences between patients with a poor and normal clinical outcome, especially within the group of classical patients. The inclusion of larger cohorts is crucial to improve statistical power and to find (clinically) relevant significant differences.

To enable the combination of data from different cohorts, the standardization of clinical outcome measures is of major importance. The implementation of the International guideline for the management of CG patients will facilitate a standardized clinical follow-up of patients, especially the neuropsychological and neurological evaluation[5]. Not only patients with potential or severe long-term complications, but all CG patients should receive follow-up as described in the guideline. This will enable the inclusion of patients that reflect the entire spectrum of clinical outcome of CG in future research. Moreover, standardized longitudinal follow-up studies could clarify if CG is a (neuro)degenerative disease. Even though many studies investigated the long-term complications of CG, standardized longitudinal research is lacking and it remains uncertain if a decline in neurological, neuropsychological and endocrinological functioning over time occurs. The lower frequency of movement disorders in children compared with adults found in our cohort is in agreement with previous findings[17] and suggests a progressive course of neurological disease. This finding underlines the need for standardized longitudinal research in CG. Although structural changes of both gray- and white matter were found on the brain MRIs of CG patients, the course of gray- and white matter changes and the timing of damage to the brain remains unclear. Longitudinal MRI research that uses quantitative neuroimaging parameters could be of use to investigate if damage to the brain occurs early on or whether CG is a neurodegenerative disorder with progressive (neuronal) damage throughout life.

In this thesis, we used a pre-defined cut-off for a poor and normal clinical outcome as defined by the presence and absence of intellectual, neurological and endocrinological complications. However, such a cut-off does not reflect the observed individual variation in clinical symptoms, which is a continuum. The impact of complications on daily functioning and the quality of life of patients were also not part of this thesis and could

be of value. Future research should aim for clear definitions that indicate the presence and extent of long-term complications, especially if CG turns out to be a neurodegenerative disorder. In addition, it is important to search for outcome measures that reliably reflect functioning of patients. IQ is frequently used as clinical outcome measure in studies, but as we demonstrated in **Chapter 4**, the IQ of patients may not reflect underlying more specific cognitive abilities and other cognitive outcome measures such as executive functioning, may provide a better reflection of performance than IQ.

The introduction of CG in the newborn screening program in 2007 has resulted in the detection of variant patients with previously unreported genotypes and phenotypes. These patients were asymptomatic at diagnosis, demonstrate erythrocyte GALT enzyme activity up to 10% and have undetectable erythrocyte Gal-1-P levels on dietary treatment. As the clinical and biochemical outcomes of variant patients differs from classical patients, it is important to identify variant patients and to distinguish them from classical patients, especially in studies investigating prognostic (bio)markers.

In the group of classical patients, clinical outcome was not associated with biochemical outcomes, and differences in Gal-1-P, IgG *N*-glycans and residual galactose oxidation capacity do not seem to explain the variability in clinical outcome. The differences in whole body galactose oxidation capacity were barely detectable within the group classical patients. This may result from the inclusion of a small cohort in which small differences in oxidized 1-^{13}C labeled galactose may not reflect clinically relevant differences in galactose oxidation capacity. Also, it could very well be that residual galactose oxidation capacity may not be causing the observed differences in clinical outcome. Therefore, other fields of interest should be explored in order to detect valid markers for clinical outcome. Research that focusses on primarily affected tissues such as the brain and ovaries in animal models could provide new insights. In Maastricht (in the Netherlands), a Zebrafish GALT knockout was developed which resembles the clinical and biochemical phenotype observed in patients[18]. The assessment of Gal-1-P, *N*-glycans profiles and GALT activity itself in different tissues might be useful and could clarify why some organs seem to be more affected than others.

The development of new techniques has enabled the analysis of 'omics', small molecules such as metabolomics, glycomics and lipidomics. The analysis of these small molecules might provide new biomarkers and improve our knowledge on the pathways involved[19]. To further unravel the pathophysiology of CG, it is crucial to understand how metabolites affect each other and to investigate alternative disposal pathways.

A recent study has demonstrated a second genetic disorder in 5% of patients with a Mendelian disease[20]. Because a second disorder may affect the clinical outcome of patients, the search for biomarkers should be performed in a cohort of CG patients without a second genetic diagnosis that could influence the clinical outcome, thus ensuring that the phenotype is the result of CG only. As it has been suggested that the variability in phenotypes of IEMs could be explained by genomic and epigenetic factors[21,22], the

role of (epi)genetic factors and modifier genes should be further investigated in CG patients. The identification of modifier genes could even facilitate the development of new therapies[23].

Potential new therapies

To improve the clinical outcome of CG patients, patients must receive treatment before irreversible damage occurs. To be able to assess the potential of new therapies, it is important to establish if damage not solely occurs prenatal. Currently, research focusses on various disease-modifying options that enable the treatment of CG and have the potential to influence the course of disease. Two promising therapies will be discussed here in more detail.

In most IEMs, the accumulation of substrates causes damage. Therefore, the reduction of substrates could be a powerful tool to reduce damage and improve patient outcome. As elevated levels of Gal-1-P are presumed to cause damage, it has been suggested that the inhibition of the formation of Gal-1-P by inhibiting galactokinase (GALK) activity may reduce galactose toxicity in GALT deficient tissues. The complete inhibition of GALK will also limit the formation of UDP-sugars which are important for the glycosylation of proteins and lipids and other important processes. Therefore, the partial inhibition of GALK without fully inhibiting the formation of substrates downstream of the GALT enzyme could result in lower Gal-1-P levels and may improve the clinical outcome of patients[24,25].

Another promising therapy is mRNA treatment. Current efforts are aimed at improving the delivery of mRNA to the target cells. In an animal model, increased mGALT protein expression and enzyme activity was achieved with the use of mRNA in a GALT knockout mice[26].

Meanwhile, the search for prognostic (bio)markers continues. Valid markers will provide a better prognostication for identified patients, facilitate the development of new therapeutic options and enable individualized treatment.

REFERENCES

1. Fridovich-Keil JL, Walter JH. Part 7: Carbohydrates, Chapter 72: Galactosemia.The Online Metabolic and Molecular Bases of Inherited Disease, OMMBID; Valle D.L., Antonarakis S, Ballabio A, Beaudet A.L., Mitchell G.A.(Eds.). McGraw Hill, New York.
2. Lai K, Elsas LJ, Wierenga KJ. Galactose toxicity in animals. *International union of biochemistry and molecular biology (IUBMB) Life*. 2009;61(11):1063-74.
3. Welling L, Boelen A, Derks TG, Schielen PC, de Vries M, Williams M, et al. Nine years of newborn screening for classical galactosemia in the Netherlands: Effectiveness of screening methods, and identification of patients with previously unreported phenotypes. *Molecular genetics and metabolism*. 2017;120(3):223-8.
4. Lai K, Langley SD, Singh RH, Dembure PP, Hjelm LN, Elsas LJ, 2nd. A prevalent mutation for galactosemia among black Americans. *Journal of pediatrics*. 1996;128(1):89-95.
5. Welling L, Bernstein LE, Berry GT, Burlina AB, Eyskens F, Gautschi M, et al. International clinical guideline for the management of classical galactosemia: diagnosis, treatment, and follow-up. *Journal of inherited metabolic disease*. 2017;40(2):171-6.
6. Guerrero NV, Singh RH, Manatunga A, Berry GT, Steiner RD, Elsas LJ, 2nd. Risk factors for premature ovarian failure in females with galactosemia. *Journal of pediatrics*. 2000;137(6):833-41.
7. Webb AL, Singh RH, Kennedy MJ, Elsas LJ. Verbal dyspraxia and galactosemia. *Pediatric research*. 2003;53(3):396-402.
8. Yuzyuk T, Viau K, Andrews A, Pasquali M, Longo N. Biochemical changes and clinical outcomes in 34 patients with classic galactosemia. *Journal of inherited metabolic disease*. 2018;41(2):197-208.
9. Maratha A, Stockmann H, Coss KP, Estela Rubio-Gozalbo M, Knerr I, Fitzgibbon M, et al. Classical galactosaemia: novel insights in IgG *N*-glycosylation and *N*-glycan biosynthesis. *European Journal of Human Genetics*. 2016;24(7):976-84.
10. Stockmann H, Adamczyk B, Hayes J, Rudd PM. Automated, high-throughput IgG-antibody glycoprofiling platform. *Analytical Chemistry*. 2013;85(18):8841-9.
11. Stockmann H, Coss KP, Rubio-Gozalbo ME, Knerr I, Fitzgibbon M, Maratha A, et al. IgG *N*-Glycosylation Galactose Incorporation Ratios for the Monitoring of Classical Galactosaemia. *JIMD reports*. 2016;27:47-53.
12. Pucic M, Knezevic A, Vidic J, Adamczyk B, Novokmet M, Polasek O, et al. High throughput isolation and glycosylation analysis of IgG-variability and heritability of the IgG glycome in three isolated human populations. *Molecular and cellular proteomics*. 2011;10(10):M111 010090.
13. Coelho AI, Bierau J, Lindhout M, Achten J, Kramer BW, Rubio-Gozalbo ME. Classic Galactosemia: Study on the Late Prenatal Development of GALT Specific Activity in a Sheep Model. *The anatomical record: advances in integrative anatomy and evolutionary biology*. 2017;300(9):1570-5.
14. Carlock G, Fischer ST, Lynch ME, Potter NL, Coles CD, Epstein MP, et al. Developmental Outcomes in Duarte Galactosemia. *Pediatrics*. 2019;143(1).

15. Knerr I, Coss KP, Kratzsch J, Crushell E, Clark A, Doran P, et al. Effects of temporary low-dose galactose supplements in children aged 5-12 y with classical galactosemia: a pilot study. *Pediatric research*. 2015;78(3):272-9.
16. Nelson MD, Jr., Wolff JA, Cross CA, Donnell GN, Kaufman FR. Galactosemia: evaluation with MR imaging. *Radiology*. 1992;184(1):255-61.
17. Kuiper A, Grunewald S, Murphy E, Coenen MA, Eggink H, Zutt R, et al. Movement disorders and nonmotor neuropsychological symptoms in children and adults with classical galactosemia. *Journal of inherited metabolic disease*. 2019;42(3):451-8.
18. Vanoevelen JM, van Erven B, Bierau J, Huang X, Berry GT, Vos R, et al. Impaired fertility and motor function in a zebrafish model for classic galactosemia. *Journal of inherited metabolic disease*. 2018;41(1):117-27.
19. Mussap M, Zaffanello M, Fanos V. Metabolomics: a challenge for detecting and monitoring inborn errors of metabolism. *Annals of translational medicine*. 2018;6(17):338.
20. Posey JE, Harel T, Liu P, Rosenfeld JA, James RA, Coban Akdemir ZH, et al. Resolution of Disease Phenotypes Resulting from Multilocus Genomic Variation. N Engl J Med. 2017;376(1):21-31.
21. Kammenga JE. The background puzzle: how identical mutations in the same gene lead to different disease symptoms. *Federation of European Biochemical Societiesis (FEBS) Journal*. 2017;284(20):3362-73.
22. Eanes WF. New views on the selection acting on genetic polymorphism in central metabolic genes. *Annals of the New York Academy of sciences*. 2017;1389(1):108-23.
23. McCabe ERB. Modifier genes: Moving from pathogenesis to therapy. *Molecular genetics and metabolism*. 2017;122(1-2):1-3.
24. Bosch AM, Bakker HD, van Gennip AH, van Kempen JV, Wanders RJ, Wijburg FA. Clinical features of galactokinase deficiency: a review of the literature. *Journal of inherited metabolic disease.* 2002;25(8):629-34.
25. Lai K, Boxer MB, Marabotti A. GALK inhibitors for classic galactosemia. *Future medicinal chemistry*. 2014;6(9):1003-15.
26. Balakrishnan B, An D, Nguyen V, DeAntonis C, Martini PGV, Lai K. Novel mRNA-Based Therapy Reduces Toxic Galactose Metabolites and Overcomes Galactose Sensitivity in a Mouse Model of Classic Galactosemia. *Molecular therapy*. 2020;28(1):304-12.

Chapter 10

Summary
Samenvatting

SUMMARY

Classical Galactosemia is an inborn error of metabolism, a disease in which galactose (milk sugar) is not processed correctly in the cells of the body. Lactose from milk and dairy products is converted into galactose and glucose, mainly in the gut. Normally, galactose is transported through the bloodstream and used by the body for energy after it has been converted into glucose in the liver. The enzyme galactose-1-phosphate uridylyltransferase (GALT) is an enzyme essential for one of the steps in the conversion of galactose into glucose. Galactosemia literally means galactose in the blood. In patients with Classical Galactosemia (CG), the enzyme GALT is not working properly as a result of inherited changes in the DNA. The accumulation of galactose in the blood leads to toxic products upstream of the GALT enzyme and a shortage of products downstream of the GALT enzyme. Both mechanism are presumed to contribute to the complications observed in patients. Patients are diagnosed with CG if the activity of the GALT enzyme measured in red blood cells is below 15% of controls and/or if two disease-causing mutations are found in the DNA.

In newborns with CG, the ingestion of galactose from breast milk or infant formula results in a life-threatening disease in the first weeks of life affecting multiple organs, such as the liver, kidneys, the eyes and the brain. The only available treatment is to strictly limit the intake of galactose from the diet. After the start of a galactose-restricted diet, newborns with CG quickly recover. However, most patients develop long-term complications affecting the brain resulting in motor developmental delay and speech developmental delay, a below average IQ, cognitive impairment and movement disorders. In most female patients, the ovaries are affected as well, leading to an early menopause, also known as premature ovarian insufficiency (POI). Even the early detection of patients by family screening (the screening of siblings of CG patients) with an immediate start of galactose restriction after birth, and the inclusion of CG in the newborn screening program (in the Netherlands since 2007) leading to a diagnosis and start of treatment in the first week of life, does not prevent long-term complications. The exact disease mechanism and timing of damage is not yet clear. The clinical outcome of patients is highly variable and ranges from severely affected to completely normal, which is poorly understood. To be able to prevent long-term complications in CG, more insight into the long-term outcome of patients is needed. Markers that are able to predict the course of disease are yet to be discovered but urgently needed as they can provide clarity for patients and families. In addition, these markers may serve as target for the development of new treatment options and could be used to assess their effect.

The aim of this thesis is to elucidate the broad spectrum of clinical outcome of CG and to search for (prognostic) markers of clinical outcome. **Part 1** of this thesis explores the clinical and biochemical outcome of patients and in **Part 2** potential (bio)markers and their association with clinical outcome are investigated. As galactose is present in many

products, a lifelong galactose-restricted diet can be a challenge for patients. Therefore, we developed patient education materials for children and adolescents with inborn errors of metabolism and tested if these materials improve knowledge on the disease and treatment in **Part 3**.

Part 1 The spectrum of clinical outcome

The clinical outcome of patients is highly variable and patients and even siblings with identical disease-causing mutations can be found on opposite ends of the clinical outcome spectrum. In **Chapter 2**, both the symptoms in the newborn period and the prevalence and extent of long-term complications were evaluated. In our cohort of 56 patients, differences in mutations (genotypes) and the severity of complications (phenotypes) were found. In our cohort, we were able to distinguish three different groups of patients. The first group of patients includes 47 patients with classical phenotypes with low to undetectable GALT enzyme activity in red blood cells (<3%) and two disease-causing mutations (hereafter named classical patients). This group consists of 32 patients who were diagnosed before 2007 after they presented with CG related symptoms (on average diagnosed and treated within the second week with a maximum of six weeks) and of 15 patients detected by family screening or newborn screening since 2007 who were treated early (within the first week of life). A majority of the classical patients have long-term complications, but the clinical outcome varies considerably within this group.

The second group of patients includes seven patients detected by newborn screening with previously unreported genotypes, residual GALT enzyme activity in red blood cells up to 10% and no symptoms at diagnosis (hereafter variant patients).

The third group includes two patients with two specific 'S135L' mutations found in patients of African descent, causing GALT deficiency in red blood cells but residual GALT enzyme activity in other tissues (hereafter patients with two S135L mutations). Both patients with two S135L mutations were diagnosed before 2007 and were diagnosed late after they presented with CG related symptoms at age 7 months (feeding difficulties and cataract) and age 10 years (visual impairment due to cataract).

To investigate if an early treatment resulted in a more favorable long-term outcome, the prevalence of long-term complications was investigated within the group of classical patients. Even though patients detected through family screening demonstrated no symptoms in the newborn period and the patients detected by newborn screening demonstrated fewer and less severe illness in the newborn period than patients diagnosed before 2007, an early dietary treatment did not result in a more favorable long-term outcome. The long-term outcome described in **Chapter 2** includes the neurological outcome of patients and the endocrinological outcome (hormonal and reproductive status) of female patients. The intellectual outcome and cognitive abilities of patients (neuropsychological outcome) are described in **Chapter 4**.

Neurological outcome

Of the 56 patients in our cohort, 36 were examined for neurological signs and symptoms. In 17 out of 36 patients (47%) a movement disorder was found, which was more frequently in adults (8 out of 14, 57%) than in children (9 out of 22, 41%). All patients with a movement disorder were classical patients. All variant patients and both patients with two S135L mutations were examined for neurological signs and symptoms, but no movement disorder was found. The movement disorders varied from a mild tremor to a severe tremor (involuntary muscle contractions leading to shaking of one or more parts of the body) and dystonia (involuntary muscle contractions leading to abnormal posture and spasms) interfering with daily functioning. In patients with a movement disorder, the frequency of an abnormal motor development (42%) and speech development (75%) was higher when compared with patients without a movement disorder in which 16% had an abnormal motor development and 38% had an abnormal speech development.

Endocrinological outcome

Data were available for 21 female patients aged 12 years and older. In 8 out of 19 female patients, a delayed puberty was induced with medication. In two patients information was missing. Primary ovarian insufficiency (POI) was diagnosed in 12 out of 17 (71%) female patients and was uncertain in four. The female variant patients were too young for a endocrinological evaluation. Both patients with two S135L mutations were not diagnosed with POI.

Neuropsychological outcome

Besides the intellectual abilities, neuropsychological functioning of CG patients was investigated both in the literature (**Chapter 3**) and in our own cohort (**Chapter 4**). Most studies focus on the IQ of patients rather than investigating the underlying and more specific abilities involved in knowledge gain and comprehension (cognition). Cognitive functioning can be divided into multiple domains such as information processing speed, attention, memory, visuospatial functioning (the ability to process and interpret visual information) and executive functioning (an overall term used for skills that include working memory, flexible thinking and self-control). As these domains provide information on the specific abilities of patients, it is important to investigate them separately. In **Chapter 3**, we reviewed the available literature on cognitive functioning in CG patients and identified 11 studies. As a group, patients had below average scores on all cognitive domains. However, the results differed between studies, most studies used a limited number of tests or reported on small cohorts. Also, most studies were of limited quality. In **Chapter 4**, we investigated multiple cognitive domains and social functioning in our cohort of patients. The 48 included patients had an average IQ of 77 (range 45-103), which is below average since the average IQ of the general population

is 100 and an IQ below 85 is below average. All variant patients had an IQ above 85, while both patients with two S135L mutations had an IQ well below 85. In the group of classical patients, a majority of the patients (68%) had an IQ below 85. Compared to the general population, CG patients demonstrated impaired cognitive functioning without a clear profile. The results indicated that cognitive outcomes of patients may exceed IQ based expectations and thus the IQ may not reflect the abilities of patients. As it has been suggested that CG patients demonstrate autistic traits, we investigated certain aspects of social functioning (*i.e.* social responsiveness). On average, the (parents of) patients reported normal levels of social responsiveness without features suggestive of autism. Most results were comparable to the general population, implying that social responsiveness is not impaired in CG. Since social problems, withdrawn behavior and anxiety were frequently reported by (parents of) patients, social functioning should be further investigated. Based on the large variability in cognitive, behavioral and social functioning without a clear profile, a neuropsychological evaluation and follow-up is advised in all patients to provide patients with support if needed.

Brain abnormalities on MRI

As the long-term complications mainly involve the brain, the brain MRIs of CG patients were evaluated in **Chapter 2**. In 86% of the patients brain abnormalities were found, most frequently a reduced brain volume (*i.e.* atrophy of the cerebrum and cerebellum). White matter lesions indicating damage to the white matter were found in patients, but the extent of visible white matter lesions was limited. The presence of white matter lesions was not associated with IQ nor with movement disorders, but atrophy of the cerebrum and cerebellum was more frequently present in patients with a lower IQ and/or a movement disorder. As changes in the brain may not be visible on MRI, computer programs were used to evaluated gray- and white matter in detail and the results are described in **Chapter 7**.

Biochemical outcome

The biochemical outcome of patients is described in **Chapter 2**. The biochemical outcome of the seven variant patients and two patients with two S135L mutations differed from the 47 classical patients.

In CG patients, the accumulation of toxic products upstream of the GALT enzyme, such as Gal-1-P and galactitol can be measured in red blood cells and urine respectively. In classical patients, both Gal-1-P and galactitol remain elevated despite restriction of galactose in the diet. In all but one variant patient, Gal-1-P levels were below the detection limit during dietary treatment just as in a healthy population and galactitol levels were within the normal range. This indicates that the variant patients, who currently demonstrate no long-term complications, biochemically differ from classical patients. Both patients with two S135L mutations demonstrated lower Gal-1-P levels

and galactitol levels when compared with classical patients, but still higher than the levels observed in the variant patients. Therefore, we investigated if differences in the clinical outcome of CG patients are caused by individual differences in the extent of galactose intoxication (*i.e.* higher Gal-1-P levels indicate more galactose intoxication) and could be associated with a poor clinical outcome.

- ***Gal-1-P***

The most recent Gal-1-P level of CG patients was not significantly different between patients with a poor clinical outcome (IQ<85, the presence of movement disorders and/or POI in females) and with a normal clinical outcome (IQ≥85, the absence of movement disorders and/or POI). In the group of classical patients, the differences in Gal-1-P levels between patients were small and were not associated with clinical outcome.

- ***Lifetime Gal-1-P***

As the most recent Gal-1-P may not represent overall galactose intoxication, lifetime Gal-1-P was evaluated as well. Gal-1-P has been reported to reach a steady state within a few months up to one year after the start of a galactose-restricted diet. Therefore, we defined lifetime Gal-1-P as the average of all available Gal-1-P levels reported after 12 months. As variant patients and patients with two S135L mutations demonstrated a different biochemical profile, lifetime Gal-1-P was only evaluated in classical patients. Just as the most recent Gal-1-P level, lifetime Gal-1-P was not significantly different between patients with and without long-term complications. To investigate if Gal-1-P had reached a steady state after 12 months, we studied its association with age. The finding of a significant negative correlation indicates a more gradual decline of Gal-1-P after 12 months. This is an important finding as Gal-1-P is still considered as a potential marker of clinical outcome.

A shortage of products downstream of the GALT enzyme leads to a reduced production of 'UDP sugars', which are needed for the production of proteins and lipids. Glycosylation is a process in which sugars are added to proteins and lipids creating special glycoproteins and glycolipids that are involved in many processes of the body. In CG patients, glycosylation abnormalities have been demonstrated. Therefore, we investigated if glycosylation abnormalities (as indicated by *N*-glycan profiles) were associated with the clinical outcome of patients.

- ***IgG N-glycans***

To be able to investigate glycosylation abnormalities, glycoproteins and more specifically *N*-linked glycans on Immunoglobulin G (IgG) were analyzed, as IgG has well defined glycan structures which have been associated with the development of diseases (e.g. autoimmune diseases, infectious diseases and cancer). Considering IgG is present in

the blood and blood samples are easy to collect, this method is minimally invasive for patients. The *N*-glycans were analyzed with the use of an automated glycan preparation method linked to ultra-performance liquid chromatography, which is used to separate and distinguish *N*-glycans from other components. Even though significant differences in *N*-glycans were found between patients and controls, *N*-glycans were not significantly different between patients with a poor and a normal intellectual, neurological and endocrinological outcome.

Part 2 Towards individual prognostication

Galactose oxidation capacity

In this thesis, we hypothesized that differences in clinical outcome are caused by differences in galactose oxidation capacity. A (slightly) higher enzyme activity might lead to a (clinically relevant) higher ability to convert galactose (*i.e.* galactose oxidation capacity), which will cause lower Gal-1-P levels and less abnormal glycosylation and possibly a more favorable clinical outcome.

In **Chapter 5**, we studied galactose oxidation capacity in cultured skin cells (*i.e.* fibroblasts) of CG patients and controls. We chose fibroblasts because performing a skin biopsy is minimally invasive and skin cells can be multiplied easily by culturing. We used a previously developed method (galactose metabolite profiling, GMP) to determine galactose oxidation capacity. The Galactose Index (GI) (*i.e.* the ratio of [U^{13}C]Gal-1-P / [$^{13}C_6$]UDP-galactose) distinguished classical patients from controls, from patients with two S135L mutations and from variant patients. As Gal-1-P decreases and UDP-galactose increases in patients with higher galactose oxidation capacity, a lower GI indicates more residual galactose oxidation capacity. The controls in our cohort demonstrated a barely detectable GI (0.84 – 1.12) and the variant patients demonstrated a GI in the low range (1.61 – 7.17). Interestingly, the patients with two S135L mutations demonstrated a lower GI than most variant patients (2.47 and 5.40), indicating more residual galactose oxidation capacity in fibroblasts. The classical patients demonstrated a large variability in GI (9.04 – 22.37) which was not associated with clinical outcome (*i.e.* a higher GI was also found in patients with a normal clinical outcome).

In **Chapter 6**, we determined whole body galactose oxidation capacity in patients and controls by using a non-invasive 1-^{13}C galactose breath test. Galactose is a sugar, which consists of six carbon atoms. For this study one out of six atoms was labeled. Breath samples were collected for patients and controls both before and 60, 90 and 120 minutes after they ingested a harmlessly small amount of labeled galactose. The amount of labeled galactose in the breath samples was measured with the use of gas-isotope-ratio mass spectrometry, which can reliably distinguish small molecules. Hereafter, the percentage of oxidized galactose (CUMPCD) was calculated. The 1-^{13}C galactose breath test was able to distinguish classical patients (median CUMPCD T120: 0.35) from variant patients (median CUMPCD T120: 13.79), from patients with two S135L

mutations (median CUMPCD T120: 9.44) and from controls (median CUMPCD T120: 9.29). In the group of variant patients with residual GALT enzyme activity in red blood cells up to 10% and undetectable Gal-1-P levels, the whole body galactose oxidation capacity was within the control range. The higher galactose oxidation capacity and lower Gal-1-P levels compared with classical patients and the finding that the variant patients currently have a normal clinical outcome supports our hypothesis. As the differences in galactose oxidation capacity within the group of classical patients were small and galactose oxidation capacity was not able to differentiate between classical patients with a poor and normal clinical outcome, the potential of residual galactose oxidation capacity as (bio)marker in patients with CG seems limited.

MRI of the brain studied in more detail

In **Chapter 7**, various MRI techniques and analysis programs were used to investigate the brain of 21 CG patients and 24 controls. Of the 21 patients, 15 patients (71%) had an IQ below 85. In total 16 out of 21 patients were examined for neurological signs and symptoms and in 9 out of 16 patients (56%) a movement disorder was found. In the brain, white matter is responsible for the communication between nerves and gray matter processes the information. On MRI, the difference between gray- and white matter is visible due to the higher content of myelin in white matter. Myelin is a mixture of proteins and lipids that form a protecting sheath around nerves. White matter abnormalities are frequently reported in CG and therefore we investigated both total white matter and a specific white matter tract, the corticospinal tract, which is responsible for controlling movements of the limbs and torso. The gray- and white matter volume, white matter integrity and myelin content in both gray matter, white matter and the CST was lower for patients when compared to controls. Moreover, these values were associated with the neurological and intellectual outcome of patients (*i.e.* gray and white matter volumes, white matter integrity and myelin content were lower for patients with more severe movement disorders and/or a lower IQ when compared to patients without long-term complications). The extent of visible white matter lesion was limited in CG patients as the white matter lesion volume of patients was comparable to healthy subjects. The results of this study demonstrate that not only white matter is affected in CG patients, but gray matter as well. The finding that the structural changes in the brain of CG patients were associated with neurological and intellectual outcome indicate that MRI can be of use to further unravel the disease mechanism of CG. The clinical outcome of patients (*i.e.* cognitive impairment and movement disorders), the absence of widespread white matter lesions and the presence of both gray- and white matter changes on MRI could indicate that CG is primarily a gray matter disease with secondary damage to the white matter. As this was an explorative study in a small patient group, more research is needed in a larger cohort in which the course of gray- and white matter should be investigated by repeating the MRI over time.

Part 3 Patient education

The final part of this thesis addresses patient education. The galactose-restricted diet is currently the only available treatment and patients are advised to stick to this diet for the rest of their lives. Galactosemia and inborn errors of metabolism in general are complex, and children and adolescents suffering from these disorders may benefit from patient education materials that improve knowledge on disease and treatment. In the study described in **Chapter 8**, patient education materials were developed for children and adolescents with galactosemia as well as other inborn errors of metabolism and were tested. As results demonstrated a significant knowledge gain in patients, these materials could be of use to explain complex diseases and treatment to children and adolescents suffering from inborn errors of metabolism.

SAMENVATTING

Klassieke galactosemie is een erfelijke stofwisselingsziekte, waarbij galactose (melksuiker) niet goed wordt verwerkt in de cellen van het lichaam. Lactose uit melk en zuivelproducten wordt in de darm omgezet in galactose en glucose. Normaal gesproken wordt galactose in de lever omgezet in glucose. Vervolgens wordt het door de bloedbaan vervoerd en door het lichaam gebruikt voor energie. Het enzym galactose-1-fosfaat uridylyltransferase (GALT) is een enzym dat verantwoordelijk is voor een belangrijke stap in de omzetting van galactose naar glucose. Galactosemie betekent letterlijk galactose in het bloed. In patiënten met Klassieke Galactosemie (KG) werkt het GALT-enzym niet goed door erfelijke veranderingen in het DNA. Hierdoor stijgt de hoeveelheid galactose in het bloed wat leidt tot zowel de ophoping van giftige producten voorafgaand aan de omzetting door het GALT-enzym als een tekort aan producten die worden gevormd tijdens de omzetting door het GALT-enzym. Er wordt verondersteld dat beide mechanismen bijdragen aan de complicaties die bij patiënten worden waargenomen. Patiënten krijgen de diagnose KG als de activiteit van het GALT-enzym in rode bloedcellen lager is dan 15% ten opzichte van controles en/of als er twee ziekteverwekkende mutaties in het DNA worden gevonden.

Bij pasgeborenen met KG leidt de inname van galactose uit moedermelk of zuigelingenvoeding in de eerste levensweken tot een levensbedreigende ziekte die meerdere organen aantast, zoals de lever, de nieren, de ogen en de hersenen. Op dit moment is de enige beschikbare behandeling het strikt beperken van galactose in de voeding. Na de start van een galactose beperkt dieet herstellen pasgeborenen met KG snel. Ondanks een galactose beperkt dieet ontwikkelen de meeste patiënten lange-termijn complicaties. Deze complicaties worden voornamelijk veroorzaakt door schade aan de hersenen met als gevolg een ontwikkelingsachterstand van zowel de motoriek als de spraak, een ondergemiddeld IQ, cognitieve stoornissen en bewegingsstoornissen. Bij de meeste vrouwelijke patiënten worden ook de eierstokken aangetast, wat leidt tot een vroege menopauze, ook wel primaire ovariële insufficiëntie (POI) genoemd. De vroege opsporing van patiënten, enerzijds door familiescreening (de screening van broers en zussen van KG patiënten) waardoor het dieet direct na de geboorte gestart kan worden en anderzijds door de toevoeging van KG aan de hielprikscreening voor pasgeborenen (in Nederland sinds 2007) en hierdoor een vroege start van het dieet in de eerste levensweek, voorkomt lange-termijn complicaties niet. Het exacte ziektemechanisme en het moment waarop schade optreedt is nog niet duidelijk. Het is op dit moment nog niet precies bekend hoe het komt dat de klinische uitkomst van patiënten erg verschillend is, waarbij sommige patiënten heel veel last hebben van complicaties en andere patiënten juist weinig tot geen last. Om complicaties bij KG te kunnen voorkomen is meer inzicht nodig in de lange-termijn uitkomst van patiënten. Markers, die het beloop van de ziekte kunnen voorspellen, zijn nog niet beschikbaar, maar zijn dringend nodig. Deze markers

kunnen meer duidelijkheid geven over het te verwachte ziektebeloop, wat erg belangrijk is voor patiënten en hun familie. Bovendien kunnen deze markers gebruikt worden om nieuwe behandelingen voor KG te ontwikkelen en om het effect van deze behandelingen te beoordelen.
Het doel van dit proefschrift is om het brede spectrum van klinische uitkomst van KG op te helderen en te zoeken naar markers die de klinische uitkomst van patiënten kunnen voorspellen. **Deel 1** van dit proefschrift onderzoekt de klinische en biochemische uitkomst van patiënten en in **Deel 2** worden potentiële (bio) markers en hun associatie met klinische uitkomst onderzocht. Galactose is aanwezig in vele producten en een levenslang galactose beperkt dieet kan om deze reden een uitdaging zijn voor patiënten. Daarom hebben wij in **Deel 3** voorlichtingsmateriaal ontwikkeld voor kinderen en adolescenten met een aangeboren stofwisselingsziekte en getest of dit de kennis over ziekte en behandeling bij patiënten kan verbeteren.

Deel 1 Het spectrum van klinische uitkomst
De klinische uitkomst van patiënten is zeer variabel en patiënten en zelfs broers en zussen met dezelfde ziekteverwekkende mutaties worden aan beiden kanten van het klinische uitkomst spectrum gevonden. In **Hoofdstuk 2** werden zowel de symptomen bij pasgeborenen als de lange-termijn complicaties onderzocht. In ons cohort van 56 patiënten werden verschillende mutaties (genotypes) gevonden en was er een duidelijk verschil in hoe de ziekte zich uit / de ernst van de complicaties (fenotypes). In ons cohort onderscheidden wij drie groepen patiënten. De eerste groep patiënten bestaat uit 47 patiënten met de klassieke vorm van KG. Zij hebben een klassieke ziektepresentatie met lage tot niet-detecteerbare GALT-enzymactiviteit in rode bloedcellen (<3%) en twee bekende ziekteverwekkende mutaties (hierna klassieke patiënten genoemd). Binnen deze groep werden 32 patiënten vóór 2007 gediagnosticeerd nadat zij zich presenteerden met KG-gerelateerde symptomen, wat leidde tot de diagnose en het starten van het dieet in de 2e tot de 6e levensweek. Daarnaast zijn er 15 klassieke patiënten die de diagnose naar aanleiding van familiescreening hebben gekregen of sinds 2007 zijn gediagnosticeerd naar aanleiding van de hielprikscreening en hierdoor zijn deze patiënten vroeg (binnen de eerste levensweek) behandeld. Een meerderheid van de patiënten met de klassieke vorm van KG heeft lange-termijn complicaties, maar de klinische uitkomst varieert sterk binnen deze groep.
De tweede groep bestaat uit zeven patiënten die allen zijn gediagnosticeerd naar aanleiding van de hielprikscreening. Zij hebben niet eerder gerapporteerde genotypes, hadden geen symptomen ten tijde van de diagnose en in rode bloedcellen werd een GALT enzym activiteit gemeten tussen de 3 en 10% (hierna variant patiënten genoemd).
De derde groep omvat twee patiënten met in tweevoud de specifieke 'S135L' mutatie, die gevonden wordt bij patiënten van Afrikaanse afkomst. Deze mutatie veroorzaakt een lage GALT enzym activiteit in rode bloedcellen (<3%), maar restactiviteit van het

GALT enzym in andere weefsels (hierna patiënten met twee S135L mutaties genoemd). Beide patiënten met twee S135L mutaties werden gediagnosticeerd vóór 2007 maar de diagnose werd laat gesteld nadat ze KG-gerelateerde symptomen vertoonden op de leeftijd van 7 maanden (voedingsproblemen en staar) en op de leeftijd van 10 jaar (verminderd zicht door staar).
Om te onderzoeken of patiënten die vroeg behandeld zijn een gunstiger klinisch beloop hebben, werden de lange-termijn complicaties onderzocht binnen de groep klassieke patiënten. Hoewel patiënten die werden opgespoord via familiescreening na de geboorte niet ziek waren en patiënten die door de hielprikscreening werden opgespoord minder ziek waren dan patiënten die vóór 2007 werden gediagnosticeerd, leidde het vroeger starten met een dieet niet tot minder lange-termijn complicaties. De lange-termijn uitkomsten die onderzocht werden in **Hoofdstuk 2** bestaan uit de neurologische uitkomst van patiënten en de endocrinologische uitkomst (hormonale en reproductieve status) van vrouwelijke patiënten. Het IQ en de cognitieve vaardigheden van patiënten (de neuropsychologische uitkomst) werden onderzocht in **Hoofdstuk 4**.

Neurologische uitkomst
Van de 56 patiënten werden 36 patiënten neurologisch onderzocht. Bij 17 van de 36 patiënten (47%) werd een bewegingsstoornis vastgesteld, vaker bij volwassenen (8/14, 57%) dan bij kinderen (9/22, 41%). Alle patiënten met een bewegingsstoornis hadden de klassieke vorm van KG. Alle variant patiënten en beide patiënten met twee S135L mutaties zijn neurologisch onderzocht, maar geen van allen had een bewegingsstoornis. De bewegingsstoornissen varieerden van een milde tremor tot een ernstige tremor (onvrijwillige spiersamentrekkingen die leiden tot het trillen van een of meerdere lichaamsdelen) en milde tot ernstige dystonie (onvrijwillige spiersamentrekkingen die leiden tot een abnormale houding en spasmen) die het dagelijks functioneren beïnvloeden. Patiënten met een bewegingsstoornis hadden vaker een motorische ontwikkelingsachterstand (42%) en spraakontwikkelingsachterstand (75%) dan patiënten zonder bewegingsstoornis waarbij 16% een motorische ontwikkelingsachterstand en 38% een spraakontwikkelingsachterstand had.

Endocrinologische uitkomst
De data van 21 vrouwelijke patiënten van 12 jaar en ouder was beschikbaar. Bij 8 van de 19 vrouwelijke patiënten werd een vertraagde puberteit vastgesteld, welke werd opgewekt met medicatie. Bij twee patiënten ontbrak informatie hierover. Primaire ovariële insufficiëntie (POI) werd gediagnosticeerd bij 12 van de 17 (71%) vrouwen en was onduidelijk bij vier patiënten. De vrouwelijke variant patiënten waren nog te jong voor een endocrinologisch onderzoek. Beide patiënten met twee S135L mutaties hadden geen POI.

Neuropsychologische uitkomst

De neuropsychologisch uitkomst van KG patiënten werd zowel in de literatuur (**Hoofdstuk 3**) als in ons eigen cohort (**Hoofdstuk 4**) onderzocht. De meeste studies richten zich op het IQ van patiënten in plaats van op de onderliggende en meer specifieke cognitieve vaardigheden, die belangrijk zijn voor kennisverwerving, inzicht en begrip. Cognitief functioneren bestaat uit meerdere domeinen, zoals informatieverwerkingssnelheid, aandacht, geheugen, ruimtelijk inzicht en executief functioneren (een overkoepelende term voor vaardigheden zoals werkgeheugen, flexibel denken en zelfbeheersing). Omdat deze domeinen informatie geven over de specifieke mogelijkheden van patiënten is het belangrijk om deze domeinen apart te onderzoeken. In **Hoofdstuk 3** hebben we in de literatuur gezocht naar studies over het cognitief functioneren bij KG patiënten en 11 studies gevonden. De studies rapporteerden dat patiënten als groep ondergemiddeld scoorden op alle cognitieve domeinen. Het was niet mogelijk om een eenduidige conclusie te trekken over de resultaten, omdat de meeste studies een beperkt aantal testen gebruikten, de resultaten tussen studies verschilden en de meeste studies onderzoek hebben gedaan in een kleine groep. Ook was de kwaliteit van de meeste studies beperkt. In **Hoofdstuk 4** hebben we meerdere cognitieve domeinen en het sociaal functioneren in onze eigen patiënten groep onderzocht. De 48 geïncludeerde patiënten hadden een ondergemiddeld IQ. Patiënten hadden een gemiddeld IQ van 77 (45-103), terwijl in de Nederlandse bevolking een IQ van 100 gemiddeld is en een IQ onder de 85 ondergemiddeld is. Alle variant patiënten hadden een IQ boven de 85, terwijl beide patiënten met twee S135L mutaties een IQ ruim onder de 85 hadden. Bij patiënten met de klassieke vorm van KG had een meerderheid (68%) een IQ onder de 85. In vergelijking met de Nederlandse bevolking lieten KG patiënten lagere scores zien op de cognitieve testen zonder een duidelijk profiel. Daarnaast werden sommige cognitieve testen door patiënten beter gemaakt dan de verwachting was op basis van het IQ, wat zou kunnen betekenen dat het IQ de mogelijkheden van patiënten niet goed weergeeft. Omdat er wordt gesuggereerd dat KG patiënten autistische kenmerken vertonen, hebben wij bepaalde aspecten van sociaal functioneren (o.a. sociale responsiviteit) onderzocht. Over het algemeen rapporteerden (ouders van) patiënten een normale mate van sociale responsiviteit zonder autistische kenmerken. Aangezien de meeste resultaten vergelijkbaar waren met de Nederlandse bevolking lijkt sociale responsiviteit niet te zijn aangedaan bij patiënten met KG. Aangezien sociale problemen, teruggetrokken gedrag en angst relatief vaak werden gerapporteerd door (ouders van) patiënten is het belangrijk om het sociaal functioneren verder te onderzoeken. Vanwege de grote variabiliteit in cognitief- en sociaal functioneren zonder duidelijk profiel wordt een neuropsychologische evaluatie en follow-up geadviseerd in alle patiënten, om patiënten die hulp nodig hebben de juiste hulp te kunnen bieden.

Hersenafwijkingen op 'Magnetic Resonance Imaging' (MRI)
Aangezien de lange-termijn complicaties voornamelijk betrekking hebben op de hersenen, hebben wij in **Hoofdstuk 2** de MRI's van de hersenen van KG patiënten onderzocht. Bij 86% van de patiënten werden hersenafwijkingen gevonden. De meest voorkomende afwijking was een verminderd hersenvolume (atrofie) van zowel de kleine als grote hersenen. Bij patiënten werden beschadigingen (laesies) gevonden in de witte stof, maar de mate waarin deze witte stof laesies voorkwamen was beperkt. De aanwezigheid van witte stof laesies was niet geassocieerd met het IQ van patiënten of de aanwezigheid van bewegingsstoornissen, maar atrofie kwam vaker voor bij patiënten met een lager IQ en/of een bewegingsstoornis. Omdat veranderingen in de hersenen niet altijd duidelijk zichtbaar zijn op een MRI werden computerprogramma's gebruikt om de grijze- en witte stof in detail te onderzoeken en de resultaten hiervan staan beschreven in **Hoofdstuk 7**.

Biochemische uitkomst
De biochemische uitkomst van patiënten werd onderzocht in **Hoofdstuk 2**. De biochemische uitkomst van de zeven variant patiënten en beide patiënten met twee S135L mutaties was anders dan van de 47 klassieke patiënten. De ophoping van giftige producten (galactose-1-fosfaat, hierna Gal-1-P en galactitol) voorafgaand aan de omzetting door het GALT-enzym kunnen worden gemeten in rode bloedcellen (Gal-1-P) en urine (galactitol) van KG patiënten. Bij klassieke patiënten blijven zowel Gal-1-P als galactitol verhoogd ondanks de beperking van galactose in het dieet. Op één patiënt na zijn de Gal-1-P waarden van alle variant patiënten tijdens het dieet onder de detectielimiet, net als bij gezonde mensen en de galactitol waarden van de variant patiënten waren ook normaal. Dit geeft aan dat de variant patiënten, die momenteel geen lange-termijn complicaties vertonen, biochemisch verschillen van de klassieke patiënten. Beide patiënten met twee S135L mutaties hebben lagere Gal-1-P en galactitol waarden in vergelijking met klassieke patiënten, maar de waarden zijn niet zo laag als de waarden van de variant patiënten. Wij hebben daarom onderzocht of verschillen in de klinische uitkomst van KG patiënten worden veroorzaakt door individuele verschillen in de mate van galactose-intoxicatie. Hogere Gal-1-P waarden duiden op meer galactose-intoxicatie en wij hebben onderzocht of dit geassocieerd is met een slechte klinische uitkomst.

- Gal-1-P
De meest recente Gal-1-P waarde van KG patiënten was niet significant verschillend tussen patiënten met een slechte klinische uitkomst (IQ <85, de aanwezigheid van bewegingsstoornissen en/of POI bij vrouwen) en een normale klinische uitkomst (IQ ≥85, de afwezigheid van bewegingsstoornissen en/of POI). In de groep klassieke patiënten waren de verschillen in Gal-1-P waarden tussen patiënten erg klein en ze

waren niet geassocieerd met klinische uitkomst.

- 'Lifetime' Gal-1-P

Aangezien de meest recente Gal-1-P waarde mogelijk geen goede reflectie is van de algehele galactose-intoxicatie werd ook de Gal-1-P waarde door de jaren heen geëvalueerd. Gal-1-P daalt direct na het starten met een galactose beperkte dieet en er werd aangenomen dat Gal-1-P een vast niveau bereikt binnen enkele maanden tot een jaar na het starten met het dieet. 'Lifetime' Gal-1-P hebben wij daarom gedefinieerd als het gemiddelde van alle beschikbare Gal-1-P waarden die na 12 maanden zijn gemeten per patiënt. Omdat variant patiënten en patiënten met twee S135L mutaties een ander biochemisch profiel vertoonden, werd 'lifetime' Gal-1-P alleen geëvalueerd in de klassieke patiënten. Net als de meest recente Gal-1-P waarde, was 'lifetime' Gal-1-P niet significant verschillend tussen patiënten met en zonder lange-termijn complicaties. Om te onderzoeken of Gal-1-P na 12 maanden een vast niveau bereikt, hebben we de associatie met leeftijd onderzocht en een duidelijke negatieve correlatie werd gevonden. Dit suggereert een geleidelijke afname van Gal-1-P na 12 maanden. Dit is een belangrijke bevinding, aangezien Gal-1-P nog steeds wordt beschouwd als een mogelijke marker van klinische uitkomst.

Een tekort aan producten na de omzetting door het GALT-enzym leidt tot een verminderde productie van 'UDP-suikers', die nodig zijn voor de aanmaak van eiwitten en lipiden. Glycosylering is een proces waarbij suikers worden toegevoegd aan eiwitten en vetten, waardoor speciale glycoproteïnen en glycolipiden ontstaan, die bij veel processen in het lichaam betrokken zijn. Bij KG patiënten zijn glycosyleringsafwijkingen aangetoond en daarom hebben wij onderzocht of *N*-glycanen (als maat voor glycosyleringsafwijkingen) geassocieerd zijn met de klinische uitkomst van patiënten.

- IgG N-glycanen

Om glycosyleringsafwijkingen te kunnen onderzoeken, werden *N*-glycanen, glycoproteïnen die gekoppeld zijn aan immunoglobuline G (IgG), geanalyseerd. IgG heeft goed gedefinieerde *N*-glycanen die in verband zijn gebracht met het ontstaan van ziekten (zoals auto-immuunziekten, infectieziekten en kanker). Er is voor deze methode gekozen omdat IgG aanwezig is in het bloed en bloed relatief makkelijk af te nemen is, waardoor dit onderzoek niet zo invasief is voor patiënten. De *N*-glycanen werden geanalyseerd door een apparaat dat in staat is om *N*-glycanen te scheiden en te onderscheiden van andere aanwezige stoffen. Hoewel de *N*-glycanen significant verschilden tussen patiënten en controles, waren *N*-glycanen niet significant verschillend tussen patiënten met een slechte en een normale intellectuele, neurologische en endocrinologische uitkomst.

Deel 2 De zoektocht naar individuele prognosticering

Galactose oxidatie capaciteit

In dit proefschrift hebben we onderzocht of het verschil in klinische uitkomst wordt veroorzaakt door verschillen in galactose oxidatie capaciteit met als hypothese: een (iets) hogere GALT enzymactiviteit resulteert in de omzetting van meer galactose (een hogere galactose oxidatie capaciteit), met als gevolg lagere Gal-1-P waarden en minder glycosyleringsafwijkingen, wat leidt tot een gunstiger klinisch beloop.

In **Hoofdstuk 5** hebben we de galactose oxidatie capaciteit in gekweekte huidcellen (fibroblasten) onderzocht bij patiënten en controles. Er werd gekozen om dit te onderzoeken uit te voeren in huidcellen omdat een huidbiopsie niet zo invasief is en huidcellen gemakkelijk kunnen worden vermenigvuldigd door de fibroblasten te kweken. Wij hebben gebruik gemaakt van een eerder ontwikkelde methode (galactose metaboliet profilering, GMP) om de galactose oxidatie capaciteit te bepalen. De Galactose Index (GI), de verhouding van [U^{13}C] Gal-1-P / [$^{13}C_6$] UDP-galactose) was in staat om klassieke patiënten van controles, van patiënten met twee S135L mutaties en van variant patiënten te onderscheiden. Aangezien Gal-1-P afneemt en UDP-galactose toeneemt bij patiënten met een hogere galactose oxidatie capaciteit, duidt een lagere GI op een hogere galactose oxidatie capaciteit. De controles hadden een nauwelijks detecteerbare GI (0.84 – 1.12) en de variant patiënten vertoonden een lage GI (1.61 – 7.17). De patiënten met twee S135L mutaties hadden een lagere GI dan de meeste variant patiënten (2.47 en 5.40), wat betekent deze patiënten meer galactose oxidatie capaciteit hebben in hun fibroblasten. Er werd een grote variabiliteit in de GI van de klassieke patiënten gevonden (9.04 – 22.37), die niet geassocieerd was met de klinische uitkomst. Een hogere GI en daarmee een lage galactose oxidatie capaciteit werd namelijk ook gevonden bij patiënten met een normale klinische uitkomst.

In **Hoofdstuk 6** hebben we de galactose oxidatie capaciteit van het hele lichaam bepaald in patiënten en controles met behulp van een niet-invasieve 1-^{13}C galactose-ademtest. Galactose is een suiker molecuul dat uit zes koolstofatomen bestaat. Voor deze studie werd één van deze zes atomen gelabeld. Vervolgens werd de uitademing verzameld van patiënten en controles voorafgaand aan de inname van een kleine, onschadelijke hoeveelheid van het gelabelde galactose en 60, 90 en 120 minuten nadat ze het gelabelde galactose hadden ingenomen. De hoeveelheid gelabeld galactose in de uitgeademde lucht werd gemeten met behulp van een apparaat dat op betrouwbare wijze zeer kleine moleculen van elkaar kan onderscheiden. Hierna werd het percentage galactose berekend dat door de patiënten en controles was verbrand (CUMPCD). De 1-^{13}C galactose-ademtest was in staat om klassieke patiënten (gemiddelde CUMPCD op T120: 0.35) te onderscheiden van variant patiënten (gemiddelde CUMPCD op T120: 13.79), van patiënten met twee S135L mutaties (gemiddelde CUMPCD op T120: 9.44) en van controles (gemiddelde CUMPCD op T120: 9.29). In de groep variant patiënten, die een GALT enzym restactiviteit van 10% in rode bloedcelen hebben en

niet-detecteerbare Gal-1-P waarden in het bloed, was de galactose oxidatie capaciteit van het hele lichaam in de controle range. De hogere galactose oxidatie capaciteit en lagere Gal-1-P waarden in vergelijking met klassieke patiënten en de bevinding dat de variant patiënten momenteel een normaal klinisch beloop laten zien, ondersteunt onze hypothese. Aangezien de verschillen in galactose oxidatie capaciteit binnen de groep klassieke patiënten klein waren en de test hierdoor niet in staat was om onderscheid te maken tussen klassieke patiënten met een slechte en normale klinische uitkomst, lijkt de galactose oxidatie test als (bio) marker beperkt bruikbaar in KG.

MRI van de hersenen in detail bestudeerd

In **Hoofdstuk 7** werden verschillende MRI-technieken en analyseprogramma's gebruikt om de hersenen van 21 KG patiënten en 24 controles te onderzoeken. Van de 21 patiënten hadden 15 patiënten (71%) een IQ onder de 85. Daarnaast zijn 16 van de 21 patiënten neurologisch onderzocht, waarbij in 9 van de 16 (56%) een bewegingsstoornis werd vastgesteld. In de hersenen zorgt witte stof voor de communicatie tussen zenuwen en de grijze stof is verantwoordelijk voor het verwerken van informatie. Op een MRI is het verschil tussen grijze en witte stof goed zichtbaar doordat in witte stof meer myeline aanwezig is. Myeline is een mengsel van eiwitten en vetten die een beschermende laag rondom de zenuwen vormen. Afwijkingen in de witte stof zijn eerder beschreven bij KG patiënten en daarom onderzochten wij in deze studie zowel de witte stof in het gehele brein als een specifiek kanaal voor witte stof, het corticospinale kanaal, dat verantwoordelijk is voor het aansturen van bewegingen van de ledematen en de romp. Het volume van de grijze en witte stof, de integriteit van de witte stof en het myeline-gehalte in zowel grijze stof, witte stof als het corticospinale kanaal waren lager voor patiënten in vergelijking met controles. Bovendien waren deze waarden geassocieerd met de neurologische en intellectuele uitkomst van patiënten. De volumes van grijze en witte stof, de integriteit van witte stof en het myeline-gehalte waren lager voor patiënten met een bewegingsstoornis en/of een lager IQ in vergelijking met patiënten zonder bewegingsstoornis en/of een hoger IQ.

De mate waarin witte stof beschadigingen (laesies) voorkwamen was beperkt bij KG patiënten, aangezien het volume van de witte stof laesies vergelijkbaar was met die van gezonde proefpersonen. De resultaten van deze studie tonen aan dat niet alleen witte stof wordt aangetast bij KG patiënten, maar ook grijze stof. De bevinding dat structurele veranderingen in de hersenen van KG patiënten geassocieerd waren met de klinische uitkomst van patiënten geeft aan dat het gebruik van MRI nuttig kan zijn om het ziektemechanisme van KG te ontrafelen. De klinische uitkomst van patiënten met cognitieve- en bewegingsstoornissen, de afwezigheid van uitgebreide witte stof laesies en de aanwezigheid van zowel grijze als witte stof veranderingen op MRI zouden erop kunnen wijzen dat KG in de eerste plaats een grijze stof ziekte is met secundaire schade aan de witte stof. Omdat dit een exploratief onderzoek was in een kleine patiëntengroep

is er meer onderzoek nodig in een groter cohort waarin het beloop van grijze- en witte stof wordt onderzocht door het MRI onderzoek door de tijd heen te herhalen.

Deel 3 Patiëntenvoorlichting

Het laatste deel van dit proefschrift gaat over patiëntenvoorlichting. Het beperken van galactose in het dieet is op dit moment de enige beschikbare behandeling en patiënten wordt geadviseerd zich de rest van hun leven aan dit dieet te houden. Galactosemie en aangeboren stofwisselingsziekten in het algemeen, zijn complex. Kinderen en adolescenten, die aan deze aandoeningen lijden zouden baat kunnen hebben bij voorlichtingsmaterialen die de kennis over ziekte en behandeling verbetert. In de studie beschreven in **Hoofdstuk 8**, werden voorlichtingsmaterialen ontwikkeld voor kinderen en adolescenten met galactosemie en andere aangeboren stofwisselingsziekten en werden deze getest. De resultaten laten zien dat er een duidelijke verbetering was in de kennis van patiënten en daarom zouden deze materialen gebruikt kunnen worden om complexe ziekten en de behandeling hiervan uit te leggen aan kinderen en adolescenten die lijden aan een aangeboren stofwisselingsziekte.

10

Addendum

List of contributing authors
List of publications
PhD portfolio
Curriculum vitae
Dankwoord

LIST OF CONTRIBUTING AUTHORS

Annet M. Bosch	Department of Pediatrics, Emma Children's Hospital, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands
Matthan W.A. Caan	Department of Radiology and Nuclear Medicine, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands & Department of Biomedical Engineering, Amsterdam University Medical Center, location AMC, Amsterdam, the Netherlands
Marc Engelen	Department of Pediatrics and Pediatric Neurology, Emma Children's Hospital, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands
Hyung L. Elfrink	Department of Clinical Chemistry, Laboratory Genetic Metabolic Diseases, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands & Core Facility Metabolomics, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands
Sacha Ferdinandusse	Department of Clinical Chemistry, Laboratory Genetic Metabolic Diseases, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands
Gert J. Geurtsen	Department of Medical Psychology, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands
Dewi van Harskamp	Department of Pediatrics, Stable Isotope Laboratory, Emma Children's Hospital, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands
Merel E. Hermans	Department of Medical Psychology, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands
Carla E.M. Hollak	Department of Internal Medicine, Division of Endocrinology and Metabolism, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands
Hidde H. Huidekoper	Department of Pediatrics, Center for Lysosomal and Metabolic Diseases, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands
Mirian C.H. Janssen	Department of Internal Medicine, Radboud University Medical Center, Nijmegen, The Netherlands.
E. Marleen Kemper	Department of Pharmacy, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands
Kent Lai	Department of Pediatrics, Division of Medical Genetics, University of Utah School of Medicine, United States of America

Janneke G. Langendonk Department of Internal Medicine, Center for Lysosomal and Metabolic Diseases, Erasmus MC, University Medical Center Rotterdam, The Netherlands

Johanna H. van der Lee Pediatric clinical Research Office, Amsterdam UMC, University of Amsterdam & Knowledge Institute of the Dutch Association of Medical Specialists, Utrecht, the Netherlands

Roisin O'Flaherty NIBRT GlycoScience Group, National Institute for Bioprocessing, Research and Training, Mount Merrion, Blackrock, Co, Dublin, Ireland

Kim J. Oostrom Psychosocial Department, Emma Children's Hospital, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands

Esmee Oussoren Department of Pediatrics, Center for Lysosomal and Metabolic Diseases, Erasmus MC, University Medical Center Rotterdam, The Netherlands

M. Estela Rubio Gozalbo Department of Pediatrics and Department of Clinical Genetics, Maastricht University Medical Center, Maastricht, The Netherlands

Stefan D. Roosendaal Department of Radiology, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands

Jos P.N. Ruiter Department of Clinical Chemistry, Laboratory Genetic Metabolic Diseases, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands

Radka Saldova NIBRT GlycoScience Group, National Institute for Bioprocessing, Research and Training, Mount Merrion, Blackrock, Co, Dublin, Ireland

Henk Schierbeek Department of Pediatrics, Stable Isotope Laboratory, Emma Children's Hospital, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands

Anouk Schrantee Department of Radiology and Nuclear Medicine, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands

Eileen P. Treacy National Centre for Inherited Metabolic Disorders, The Mater Misericordiae University Hospital Dublin, Ireland

Maaike C. de Vries Department of Pediatrics, Radboud University Medical Center, Nijmegen, The Netherlands

Michel van Weeghel Department of Clinical Chemistry, Laboratory Genetic Metabolic Diseases, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands & Core Facility Metabolomics, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands

Frits A. Wijburg Department of Pediatrics, Emma Children's Hospital, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands

LIST OF PUBLICATIONS

THIS THESIS

Welsink-Karssies MM, Ferdinandusse S, Geurtsen GG, Hollak CEM, Huidekoper HH, Janssen MCH, Langendonk JG, van der Lee JH, O'Flaherty R, Oostrom KJ, Roosendaal SD, Rubio-Gozalbo ME, Saldova R, Treacy EP, Vaz FM, de Vries MC, Engelen M, Bosch AM. Deep phenotyping classical galactosemia: clinical outcomes and biochemical markers. *Brain Communications*. 2020;2(1).

Hermans ME, Welsink-Karssies MM, Bosch AM, Oostrom KJ, Geurtsen GJ. Cognitive functioning in patients with classical galactosemia: a systematic review. *Orphanet journal of rare diseases*. 2019 Oct 18;14(1):226.

Welsink-Karssies MM, Oostrom KJ, Hermans ME, Hollak CEM, Janssen MCH, Langendonk JG, Oussoren E, Gozalbo MER, de Vries M, Geurtsen GJ, Bosch AM. Classical galactosemia: neuropsychological and psychosocial functioning beyond intellectual abilities. *Orphanet journal of rare diseases*. 2020 Feb 7;15(1):42.

Welsink-Karssies MM, van Weeghel M, Hollak CEM, Elfrink HL, Janssen MCH, Lai K, Langendonk JG, Oussoren E, Ruiter JPN, Treacy EP, de Vries M, Ferdinandusse S, Bosch AM. The Galactose Index measured in fibroblasts of GALT deficient patients distinguishes variant patients detected by newborn screening from patients with classical phenotypes. *Molecular genetics and metabolism*. 2020 Mar;129(3):171-176.

Welsink-Karssies MM, van Harskamp D, Ferdinandusse S, Hollak CEM, Huidekoper HH, Janssen MCH, Kemper EM, Langendonk JG, Rubio-Gozalbo ME, de Vries MC, Wijburg FA, Schierbeek H, Bosch AM. The 1-^{13}C galactose breath test in GALT deficient patients distinguishes NBS-detected variant patients but does not predict outcome in classical phenotypes. *Journal of inherited metabolic disease*. 2020 May;43(3):507-517.

Zeltner NA, Welsink-Karssies MM, Landolt MA, Bosshard-Bullinger D, Keller F, Bosch AM, Groenendijk M, Grünert SC, Karall D, Rettenbacher B, Scholl-Bürgi S, Baumgartner MR, Huemer M. Reducing complexity: explaining inborn errors of metabolism and their treatment to children and adolescents. *Orphanet journal of rare diseases*. 2019 Nov 8;14(1):248.

OTHER PUBLICATIONS

Rubio-Gozalbo ME, Hasković M, Bosch AM, Burnyte B, Coelho AI, Cassiman D, Couce ML, Dawson C, Demirbas D, Derks T, Eyskens F, Forga MT, Grunewald S, Häberle J, Hochuli M, Hubert A, Huidekoper HH, Janeiro P, Kotzka J, Knerr I, Labrune P, Landau YE, Langendonk JG, Möslinger D, Müller-Wieland D, Murphy E, Őunap K, Ramadza D, Rivera IA, Scholl-Buergi S, Stepien KM, Thijs A, Tran C, Vara R, Visser G, Vos R, de Vries M, Waisbren SE, Welsink-Karssies MM, Wortmann SB, Gautschi M, Treacy EP, Berry GT. The natural history of classic galactosemia: lessons from the GalNet registry. *Orphanet journal of rare diseases.* 2019 Apr 27;14(1):86.

Lubout CMA, Blanco FA, Bartosiewicz K, Feillet F, Gizewska M, Hollak C, van der Lee JH, Maillot F, Stepien KM, Wagenmakers MAEM, Welsink-Karssies MM, van Spronsen FJ, Bosch AM. Bone mineral density is within normal range in most adult PKU patients. *Journal of inherited metabolic disease*. 2020 Mar;43(2):251-258.

van Gulik EC, Hemke R, Welsink-Karssies MM, Schonenberg-Meinema D, Dolman KM, Barendregt AM, Nusman CM, Maas M, Kuijpers TW, van den Berg JM. Normal MRI findings of the knee in patients with clinically active juvenile idiopathic arthritis. *European journal of radiology*. 2018 May;102:36-40.

van Gulik EC, Welsink-Karssies MM, van den Berg JM, Schonenberg-Meinema D, Dolman KM, Barendregt AM, Nusman CM, Maas M, Kuijpers TW, Hemke R. Juvenile idiopathic arthritis: magnetic resonance imaging of the clinically unaffected knee. *Pediatric radiology*. 2018 Mar;48(3):333-340.

Welsink-Karssies MM, Polderman JAW, Nieveen van Dijkum EJ, Preckel B, Schlack WS, Visser G, Hollak CE, Hermanides J. Very Long-Chain Acyl-Coenzyme A Dehydrogenase Deficiency and Perioperative Management in Adult Patients. *JIMD reports*. 2017;34:49-54.

PHD PORTFOLIO

Name: Mendy M. Welsink-Karssies

PhD period: September 2017 – September 2020

Promotor: prof. dr. F.A. Wijburg

Co-promotor: dr. A.M. Bosch

Department: Pediatric metabolic diseases

	Year	Workload (ECTS)
1. General courses		
BROK ('Basiscursus Regelgeving Klinisch Onderzoek')	2017	1.0
Practical Biostatistics	2018	1.4
2. Specific courses		
Scientific Writing	2018	1.5
Data analysis in Matlab	2018	0.7
Project & time management	2019	0.6
3. Presentations		
Amsterdam Kindersymposium (oral)	2019	0.5
ESN, Utrecht (oral)	2019	0.5
Poster presentation SSIEM, Rotterdam (two posters)	2019	1.5
Bijeenkomst Galactosemie Vereniging Nederland (oral)	2019	0.5
4. (Inter)national conferences		
EGS, Amsterdam	2017	0.3
Bijeenkomst Galactosemie Vereniging Nederland	2017	0.3
Amsterdam Kindersymposium	2018	0.3
Galactosemia Foundation Conference, Denver, USA	2018	1.0
EGS, Amsterdam	2018	0.3
ESN, Utrecht	2018	0.3
Amsterdam Kindersymposium	2019	0.3
SSIEM, Rotterdam	2019	0.5
5. Lecturing		
Consultation education	2018	0.2
Pre-IHK teaching	2018	0.2
Consultation education	2019	0.3
Pre-IHK teaching	2019	0.3
6. Tutoring, Mentoring		
Mentoring bachelor thesis student	2018	0.3
Mentoring master thesis student	2018	1.0

CURRICULUM VITAE

Mendy Marcella Welsink-Karssies werd op 12 april 1990 geboren in Amsterdam. Het grootste gedeelte van haar jeugd woonde zij in Amsterdam. Haar middelbare school diploma behaalde zij aan het Hervormd Lyceum Zuid in Amsterdam in 2008. Na aanvankelijk uitgeloot te zijn in 2008, startte zij in 2009 de studie Geneeskunde aan de Universiteit van Amsterdam. Tijdens haar coschap op de afdeling anesthesiologie kreeg zij de mogelijkheid om mee te werken aan wetenschappelijk onderzoek en dit resulteerde in haar eerste publicatie. Na haar coschappen heeft zij haar semi-arts stage kindergeneeskunde en wetenschappelijke stage beiden in het Emma Kinderziekenhuis in het Amsterdam UMC gedaan. Na het behalen van het artsexamen is zij in 2016 gestart als arts-assistent op de kinderafdeling in het OLVG, locatie West. In 2017 is zij onder leiding van prof. dr. Frits Wijburg en dr. Annet Bosch begonnen aan het onderzoek naar galactosemie op de afdeling metabole ziekten in het Emma Kinderziekenhuis in het Amsterdam UMC, locatie AMC. De wekelijkse begeleiding van Annet en de samenwerking met vele anderen heeft geresulteerd in dit proefschrift. Op dit moment werkt Mendy als ANIOS Kindergeneeskunde (Arts Niet In Opleiding tot Specialist) in het Amsterdam UMC. Haar doel is om kinderarts te worden en het werk in de kliniek te combineren met wetenschappelijk onderzoek.

DANKWOORD

Terugkijkend op mijn promotietraject wil ik graag eindigen met het bedanken van iedereen met wie ik de afgelopen jaren heb samengewerkt. Zonder jullie was dit proefschrift er niet geweest en ik heb veel van jullie geleerd. Daarnaast wil ik ook mijn familie en vrienden bedanken voor de nodige afleiding buiten werk om, jullie interesse in het onderzoek en jullie geduld, vooral als ik de vraag 'hoe gaat het met je onderzoek' niet kort kon beantwoorden. De afgelopen jaren heb ik regelmatig te horen gekregen dat de geschreven tekst korter mocht, maar in dit dankwoord wil ik graag een aantal mensen uitgebreid bedanken.

Allereerst wil ik **alle patiënten** (en hun ouders) bedanken, die vrijwillig aan dit onderzoek hebben meegedaan. Zonder jullie was dit proefschrift er niet geweest. Bedankt voor de tijd die jullie hebben genomen om mee te doen, zelfs aan meerdere onderzoeken op één dag en de flexibiliteit bij het plannen van de onderzoeken. De dagen waarop ik de testen kon afnemen waren een welkome afwisseling tussen alle analyses achter de computer. Het contact met alle patiënten en hun ouders heb ik als erg bijzonder ervaren. Bedankt voor jullie vertrouwen in het onderzoek en bereidheid om naar het AMC te komen, ook al was dit voor sommigen ver weg.

Graag wil ik ook de **Galactosemie Vereniging Nederland (GVN)** en alle personen die zich voor de GVN inzetten bedanken. Het is heel bijzonder om elk jaar weer te horen dat er met veel inzet geld is opgehaald voor het onderzoek naar galactosemie. Ook hebben jullie mij geholpen met het verspreiden van informatie over het onderzoek naar patiënten toe, wat zeker heeft bijgedragen aan het aantal deelnemende patiënten. Vanaf de eerste bijeenkomst voelde ik mij meteen erg welkom.

Dr. A.M. Bosch, **Annet**, in april 2017 spraken wij samen over de mogelijkheid om onderzoek te doen. Tijdens het gesprek bleek dat er al een plan klaar lag en dat er goed was nagedacht over de verschillende projecten, wat mij gelijk aansprak. Niet veel later in september 2017 kon ik dan ook een vliegende start maken. Nogmaals dank voor de kans om aan dit project te werken, wat heb ik veel van jou geleerd in de jaren die volgden. Ondanks de drukte in de kliniek en al jouw andere taken heb jij altijd tijd voor mij vrij gemaakt. Zonder jouw begeleiding was ik nooit (zo snel) op dit punt beland. Daarnaast wil ik je bedanken voor je bemoedigende woorden en jouw vertrouwen in mij. Het was een uitdaging om de onderzoeken tegelijk op te starten en uit te voeren, maar jij bleef het positief in zien en dat werkte aanstekelijk. Ik heb veel bewondering voor wie je bent als kinderarts en onderzoeker en ik heb veel van je geleerd zowel in de spreekkamer als erbuiten.

Prof. Dr. F.A. Wijburg, **Frits**, bedankt voor jouw begeleiding en het meedenken tijdens dit traject. Jouw kritische vragen hebben mij aan het denken gezet en jouw jarenlange ervaring hebben mijn proefschrift verder aangescherpt. Veel dank voor jouw suggesties en vlotte revisies.

Geachte leden van de promotiecommissie**, prof. dr. J.B van Goudoever, dr. M. Langeveld, prof. dr. R.H.L. Houtkooper, prof. dr. J. Oosterlaan, prof. dr. M.S. van der Knaap,** hartelijk dank voor het beoordelen van mijn proefschrift. Ik kijk ernaar uit om met u van gedachten te wisselen over de inhoud van mijn proefschrift. Dear **dr. E. Murphy**, thank you for the assessment of my thesis and for attending my defense.

Collega's uit het OLVG West, bedankt voor de leerzame tijd als ANIOS en dank dat ik af en toe nog diensten mocht komen doen tijdens mijn promotietraject, dit was een welkome afwisseling. Daarnaast wil ik Dr. **Koert Dolman** in het bijzonder bedanken. Het leek jou een goed idee om te gaan promoveren en ook al was ik daar in het begin niet van overtuigd, wat heb ik veel geleerd. Zonder jou had ik niet van dit project geweten en had ik vrijwel zeker deze promotieplek niet gekregen.

Kim Oostrom en **Gert Geurtsen**, bedankt voor jullie inzet, tijd en het delen van jullie expertise. Jullie hebben mij een kijkje gegeven in de complexe wereld van de neuropsychologie, die door jullie input toch wat inzichtelijker werd. De aanhouder wint bleek ook maar weer voor het neuropsychologie stuk. Veel dank voor jullie geduld en input.

Henk Schierbeek en **Dewi van Harskamp**, bedankt voor jullie hulp met de stabiele isotopen studie. Deze ingewikkelde materie is door jullie een stuk begrijpelijker geworden.

Collega's van het lab GMZ en in het bijzonder **Sacha Ferdinandusse, Michel van Weeghel, Jos de Ruiter** en **Fred Vaz,** bedankt voor jullie hulp zowel bij de fibroblasten studie als daarbuiten. Ik heb veel van jullie geleerd.

Anouk Schrantee, wat een klus de MRI's. Ik wist dat dit buiten mijn expertise lag, maar ik had werkelijk waar geen idee waar ik aan begon. We zijn tegen veel dingen aan gelopen, maar wat heb ik veel geleerd van dit project. Zowel van het technische aspect als van het herhaaldelijk opnieuw beginnen. Dat was niet altijd makkelijk, maar wel noodzakelijk en ik heb heel veel aan jouw begeleiding gehad. Jij was ondanks je drukke schema erg flexibel in het plannen van afspraken en het analyseren van de data, wat erg fijn werkte.

Stefan Roosendaal, bedankt voor jouw hulp bij het beoordelen van de MRI's. Dankzij jouw expertise hebben we in korte tijd zowel de kwalitatieve als kwantitatieve MRI-data kunnen verwerken en het was fijn om iemand met jouw neuro-radiologische kennis en kritische blik in ons team te hebben.

Sandra van den Berg en **Raschel van Luijk**, bedankt voor jullie hulp bij het maken van de MRI-scans. In zeer korte tijd hebben jullie het mogelijk gemaakt om zelfs in het weekend de allerjongste en angstige patiënten uit ons cohort te scannen. Zonder jullie was het MRI-onderzoek niet mogelijk geweest.

Marc Engelen, bedankt voor de fijne samenwerking en jouw hulp bij het neurologisch in kaart brengen van de patiënten en je flexibiliteit hierbij. Als ik vragen had kon ik altijd even langslopen, je maakte altijd tijd ook al had je het erg druk en dat maakte alles een stuk makkelijker. Naast je neurologische expertise was jij ook mijn peer support, maar we kwamen er al snel achter dat deze functie tijdens mijn project wellicht wat overbodig was. Toch bedankt voor jouw luisterende oor en hulp bij de manuscripten.

Metabole collega's: Carla Hollak, Mirjam Langeveld, Abel Thijs en alle overige kinderartsen en internisten uit de andere centra, bedankt voor jullie input, expertise en flexibiliteit. Dankzij jullie kon ik gemakkelijk met patiënten in contact komen en konden in korte tijd vele patiënten deelnemen aan onderzoek. Zonder jullie hadden wij deze patiënten aantallen niet gehaald. Ik heb veel geleerd van jullie inhoudelijke feedback op de manuscripten. Bedankt voor jullie tijd en bijdrage aan dit proefschrift. **Esther Schneider, Evelien Tump en Lydia Veerman**, bedankt voor jullie hulp bij de planning van de patiënten, het verwerken van de gegevens en een helpende hand bij het afnemen van de huidbiopten. De fijne samenwerking met jullie heeft mijn promotietraject een stuk makkelijker gemaakt.

Gal-Net, The Galactosemia Network, an international network is of utmost importance to improve treatment, education and research in rare diseases such as Classical Galactosemia. Dear **(inter)national colleagues**, thank you for the collaboration, for the informative meetings, for sharing your knowledge and for contributing to this thesis. It was great working with you.

Hanneke van der Lee, bedankt voor jouw bijdrage aan dit proefschrift. Onderzoek doen in een kleine patiëntengroep was een uitdaging. Ik heb veel van jou geleerd waarvoor dank.

Collega's uit het Emma Kinderziekenhuis, wat is het fijn om weer terug te zijn in de kliniek. Veel dank voor het warme welkom. Dankzij jullie hulp en gezelligheid was het een goede start.

Lieve **Anne**, jaren geleden stond jij naast mij op een belangrijke dag en op mijn promotie sta je weer naast me. Ik ken weinig mensen die zo energiek en vrolijk zijn ondanks een erg drukke baan zoveel km hiervandaan. Je wordt een fantastische chirurg en ik ben blij dat je erbij bent op m'n promotie.

Lieve **Elselijn**, je bent een van de belangrijkste personen in mijn leven. Ik kan alles maar dan ook echt alles met je bespreken. Borrels, etentjes en weekendjes weg met z'n vieren zijn altijd een succes en de gesprekken met jou en **Edwin** zijn altijd verhelderend. De afgelopen tijd is er veel veranderd en het mooie is dat dit geen invloed lijkt te hebben op onze vriendschap, integendeel. Ik ben dan ook heel blij dat jullie erbij zijn op mijn promotie.

Metabollies: Gé-Ann, Stephanie, Laura, als iemand weet hoe het is om te promoveren op een onderwerp waar elke patiënt er één is dan zijn jullie het wel. Jullie hebben mijn promotietraject gemaakt tot wat het was. Alle ups en soms een kleine down, maar bovenal de gezelligheid. Het laatste deel van mijn promotie moest ik het helaas zonder jullie doen. Hopelijk zien we elkaar snel weer.

Kamergenoten H7/H8, zonder jullie hadden de afgelopen jaren er heel anders uitgezien. Dankzij de vele koffietjes, gezellige borrels, de luxe lunches, kerstdiners en alle leuke dingen die we hebben gedaan waren de afgelopen jaren een feestje. Helaas was na mijn verlof alles anders en heb ik het laatste gedeelte alleen op mijn zolderkamer moeten doen, wat nogal een contrast was met de overvolle kamer op H8. Ik heb gemerkt hoe belangrijk het is om collega's te hebben met wie je af en toe kan sparren en die begrijpen hoe frustrerend het kan zijn als alles anders loopt. Bedankt voor de gezelligheid en de support.

Lindsey, mijn voorganger, bedankt voor jouw hulp bij het opstarten van mijn promotietraject en het voorbeeld dat je bent geweest met jouw onderzoek naar galactosemie. Ik heb jouw proefschrift er nog vaak bij gepakt.

Lieve **LGD'ers**, ruim 11 jaar geleden begonnen we aan onze studie en wat is er in de tussentijd veel veranderd. Iedereen is een andere kant op gegaan, maar hoe handig is het dat de groep straks bestaat uit een psychiater, chirurg, specialist ouderengeneeskunde, fysiotherapeut, huisarts, plastisch chirurg en intensivist (ja, ook dat laatste gaat helemaal goed komen). Marieke en Lotte, lekker koffietjes doen op het AMC en de 'PHD Life' doornemen, ik ga het missen. Inmiddels zijn jullie allebei gepromoveerd en wat waren jullie een goed voorbeeld. Dakota en Annick, ook de 'Haarlem crew' is weer nieuw leven ingeblazen, op nog vele etentjes en wijntjes.

Lieve **vrienden**, de afgelopen jaren hebben jullie regelmatig naar de verhalen over mijn promotietraject geluisterd. Bedankt voor jullie interesse in mijn werk en jullie pogingen om te begrijpen waar ik mij mee bezig hield, ook al kon ik dit niet altijd begrijpelijk uitleggen. Mijn promotie heeft ervoor gezorgd dat ik niets hoefde te missen en mijn sociale leven weer helemaal op kon halen. Bedankt voor de gezelligheid en dat jullie er voor mij zijn.

Lieve **Opa en Oma Karssies en Oma Stel,** wat ben ik blij dat jullie mogen meemaken dat ik mijn promotietraject heb afgerond. Na alle vragen over mijn onderzoek en hoe het ervoor staat, kan ik niet wachten om dit proefschrift aan u te laten zien.

Lieve **Wim en Elly,** ik bof maar met jullie als schoonouders. Schoonfamilie kies je niet, maar als ik mocht kiezen zou ik zeker jullie kiezen. Hoe jullie altijd voor ons klaar staan is heel bijzonder, we zijn zo blij met jullie! Dankzij jullie hulp en de zorg voor Jens is mijn proefschrift geworden wat het moest zijn zonder concessies te hoeven doen, heel veel dank!

Lieve **pap en mam**, bedankt dat jullie altijd voor mij klaar staan. Het is fijn om te weten dat jullie er voor mij zijn als dat nodig is. Zonder jullie support had ik niet gestaan waar ik nu sta. **Esmeralda**, lieve Es, tijdens het maken van mijn proefschrift was het weer duidelijk, de creativiteit die jij hebt ontbreekt bij mij. Ondanks dat we allebei andere wegen bewandelen ben ik blij dat die wegen elkaar regelmatig kruisen en we elkaars hulp in kunnen schakelen wanneer dit nodig is.

Melvin, de laatste maar zeker niet de minste. Wat heb ik veel aan jou gehad de afgelopen jaren en wat had ik zonder jou gemoeten?! Ik kan altijd alles bij je kwijt. Je weet altijd alles te relativeren, maar dan ook echt alles waar ik soms juist het tegenovergestelde doe (en inderdaad dit helpt helemaal niet). Jouw steun en het samen sparren zijn mijn promotietraject zeker ten goede gekomen en ik heb veel aan jou gehad. Je zult ongetwijfeld blij zijn dat het erop zit. Bedankt dat je er altijd voor mij bent. Je bent een fantastische vader voor Jens. Door jullie weet ik wat echt belangrijk is en ik geniet enorm van de tijd die we hebben met z'n drieën.