Accuracy of chitotriosidase activity and CCL18 concentration in assessing type I Gaucher disease severity. A systematic review with meta-analysis of individual participant data

Tatiana Raskovalova,1 Patrick B. Deegan,2 Pramod K. Mistry,3 Elena Pavlova,2 Ruby Yang,4 Ari Zimran,4 Juliette Berger,5,6 Céline Bourgne,5,6 Bruno Pereira,7 José Labarère8,9 and Marc G. Berger5,6,10

1Laboratoire d’immunologie, Grenoble University Hospital, Université Grenoble Alpes, Grenoble, France; 2Lysosomal Disorders Unit, Department of Medicine, University of Cambridge, Addenbrooke’s Hospital, Cambridge, UK; 3Pediatric Gastroenterology and Hepatology, Yale University School of Medicine, New Haven, CT, USA; 4Gaucher Clinic, Shaare Zedek Medical Center, Hebrew University-Hadassah Medical School, Jerusalem, Israel; 5CHU Clermont-Ferrand, Hôpital Estaing, Hématologie Biologique, Clermont-Ferrand, France; 6Université Clermont Auvergne, EA 7453 CHELTER, Clermont-Ferrand, France; 7DRCI, CHU Clermont-Ferrand, Clermont-Ferrand, France; 8Quality of Care Unit, INSERM CIC1406, Grenoble University Hospital, Université Grenoble Alpes, Grenoble, France; 9TIMC-IMAG, UMR 5525 CNRS, Université Grenoble Alpes, Grenoble, France and 10CHU Clermont-Ferrand, Service d’Hématologie Clinique Adulte et Thérapie Cellulaire, Hôpital Estaing, Clermont-Ferrand, France

ABSTRACT

Chitotriosidase activity and CCL18 concentration are interchangeably used for monitoring Gaucher disease (GD) activity, together with clinical assessment. However, comparative studies of these two biomarkers are scarce and of limited sample size. The aim of this systematic review with meta-analysis of individual participant data (IPD) was to compare the accuracy of chitotriosidase activity and CCL18 concentration for assessing type I GD severity. We identified cross-sectional and prospective cohort studies by searching Medline, EMBASE, and CENTRAL from January 1995 to June 2017, and by contacting research groups. The primary outcome was a composite of liver volume >1.25 multiple of normal (MN), spleen volume >5 MN, hemoglobin concentration <11 g/dL, and platelet count <100x10^9/L. Overall, IPD included 1,109 observations from 334 patients enrolled in nine primary studies, after excluding 111 patients with undocumented values and 18 patients with deficient chitotriosidase activity. IPD were unavailable for 14 eligible primary studies. The primary outcome was associated with a 5.3-fold (95% Confidence Interval [CI]: 4.2-6.6) and 3.0-fold (95% CI: 2.6-3.6) increase of the geometric mean for chitotriosidase activity and CCL18 concentration, respectively. The corresponding areas under the receiver operating characteristics curves were 0.82 and 0.84 (summary difference, 0.02, 95% CI: -0.02 to 0.05). The addition of chitotriosidase activity did not improve the accuracy of the CCL18 concentration. Estimates remained robust in the sensitivity analysis and consistent across subgroups. Neither the chitotriosidase activity nor the CCL18 concentration varied significantly according to a recent history of bone events among 97 patients. In conclusion, the CCL18 concentration is as accurate as chitotriosidase activity in assessing hematological and visceral parameters of GD severity and can be measured in all GD patients. This meta-analysis supports the use of CCL18 rather than chitotriosidase activity for monitoring GD activity in routine practice.

Correspondence:

José Labarère
jlabarere@chu-grenoble.fr

Marc G. Berger
mberger@chu-clermontferrand.fr

Received: August 21, 2019.
Accepted: January 20, 2020.

https://doi.org/10.3324/haematol.2019.236083

©2021 Ferrata Storti Foundation

Material published in Haematologica is covered by copyright. All rights are reserved to the Ferrata Storti Foundation. Use of published material is allowed under the following terms and conditions: https://creativecommons.org/licenses/by-nc/4.0/legalcode. Copies of published material are allowed for personal or internal use. Sharing published material for non-commercial purposes is subject to the following conditions: https://creativecommons.org/licenses/by-nc/4.0/legalcode, sect. 3. Reproducing and sharing published material for commercial purposes is not allowed without permission in writing from the publisher.
Introduction

Gaucher disease (GD; OMIM#230800) is a recessively inherited lysosomal storage disorder caused by biallelic mutations in the GBA1 gene that encodes the lysosomal acid β-glucosidase. The metabolic defect in GD results in bone marrow and visceral organ infiltration by Gaucher cells (i.e., glucosylceramide-laden macrophages), leading to anemia, thrombocytopenia, hepatosplenomegaly, and skeletal manifestations. Three main GD types (I, II, and III) have been described, on the basis of the clinical features and age of onset. Type I (i.e., non-neuropathic) GD is predominant (85-94%).

GD natural course is rather unpredictable, even within subgroups of patients with the same GBA1 mutation. Therefore, evaluating the disease severity and prognosis is challenging for clinicians. In this context, a surrogate blood biochemical marker is highly desirable for assessing GD severity and helping decision-making for specific therapy initiation or adjustment. Several biomarkers have been identified, including chitotriosidase, CCL18, and glucosylsphingosine.

Chitotriosidase, the human analogue of the non-vertebrate chitinase, is directly secreted by Gaucher cells and is considered an indicator of the overall Gaucher cell burden. In patients with GD, chitotriosidase activity is about 1,000-fold higher than the normal values, and its concentration is 10- to 50-fold higher from single-center studies of relatively limited sample size. The lack of large scale, multicenter, head-to-head comparison studies have hindered full validation of these biomarkers and formulation of context of use in the clinical practice. We hypothesized that evidence on the comparative accuracy of these two biomarkers could be strengthened by the secondary analysis of individual participant data (IPD) from primary studies on patients with GD. Therefore, our primary objective was to compare the accuracy of CCL18 concentration and chitotriosidase activity for assessing hematological and visceral parameters of type I GD severity. The secondary objective was to compare the accuracy of these two biomarkers for discriminating type I GD patients with symptomatic bone events.

Methods

This systematic review with IPD meta-analysis was performed according to current guidelines and complied with the Preferred Reporting Items for Systematic review and Meta-Analysis (PRISMA)-IPD statement. The rationale and methods were pre-specified and reported in a protocol registered at PROSPERO (CRD42015027243). This meta-analysis was carried out on data from primary studies for which ethical approval had been obtained by the investigators. The Comité de Protection des Personnes Sud Est 6, Clermont-Ferrand, France (IRB 00006256) reviewed the protocol and considered that it did not qualify for biomedical research requiring patient informed consent, provided that no supplementary data would be collected from the participants enrolled in primary studies.

Eligibility criteria

Eligible studies included cross-sectional and cohort studies that measured both chitotriosidase activity and CCL18 concentration at baseline and/or at follow-up in consecutive patients with type I GD (Online Supplementary Materials and Methods). Studies with fewer than 10 participants were excluded from this systematic review.

Information sources

Studies were identified by searching Medline, EMBASE, and Cochrane Central Register of Controlled Trials (CENTRAL) from January 1995 to June 2017. Our electronic search was supplemented by scanning the reference lists of the retrieved articles and by contacting research groups.

Search strategy

The search concepts included chitotriosidase activity, CCL18, biological markers, ERT, and GD (Online Supplementary Appendices S1-3). No restriction of document type and language was applied, and no methodology filter was used.

Study selection

Two review authors (TR and JL) assessed potentially relevant full-text articles against pre-specified eligibility criteria.

Data collection

Information on primary studies were collected using a standardized data extraction form. Where possible, IPD were extracted from published articles. Otherwise, the corresponding authors or principal investigators were invited to collaborate in this systematic review project by supplying de-identified IPD. The requested IPD included baseline characteristics and pre-specified outcomes.

Organ volumes were expressed as multiples of normal (MN) adjusted for body weight.

Outcomes

The primary outcome was a composite of hemoglobin concentration <11 g/dL (<10 g/dL for patients aged 12-59 months of age),
platelet count <100x10^9/L, spleen volume >5 MN, and liver volume >1.25 MN. The secondary outcomes included symptomatic bone events, a composite of hemoglobin concentration <8 g/dL (<7 g/dL for patients 12-59 months of age), platelet count <50x10^9/L, spleen volume >15 MN, and liver volume >2.5 MN, and individual components of the primary and secondary composite outcomes. Bone events included skeletal fracture, osteonecrosis, or avascular necrosis that occurred within the previous 12 months of biomarker analysis. All outcomes and cut-off values for continuous parameters were set according to published guidelines or previous studies, and were pre-specified.

**Statistical analysis**

We performed logarithm transformation for chitotriosidase.

---

**Figure 1. Identification and selection of studies for the meta-analysis of individual participant data.** *Seventy three additional participants (222 observations) had splenectomy at baseline. IDP: individual participant data.*
activity and CCL18 concentration and derived geometric mean ratios along with 95% Confidence Intervals (CI), in order to account for skewed distributions.24,25 The comparative accuracy of CCL18 concentration relative to chitotriosidase activity in discriminating patients with the outcomes of interest was quantified by the difference in the areas under the receiver operating characteristic (AUC-ROC) curves. Data synthesis was performed with one- and two-stage approaches.26,27

Results

Study selection and obtained IPD

Overall, 2,636 records were identified by database searching (Figure 1). Two additional records were found by contacting field experts and searching clinical trial registries, respectively. After removing 18 duplicates, the titles and abstracts of 2,625 records were screened for eligibility. Of these, 318 records were identified as being potentially relevant and full-text articles were retrieved for a more thorough review. After excluding 295 records based on the full-text article, our systematic review included 23 primary studies from which IPD were sought. IPD were available only for nine of these studies totaling 463 patients with type I GD (Figure 1). The sponsor of four of trials of ERT with velaglucerase17,31 and the principal investigators of three observational studies17,32,33 supplied computerized IPD upon request. The IPD of the other two randomized controlled trials on ERT with taliglucerase were extracted from the published article34 (Online Supplementary Appendix S4). Our meta-analytical sample consisted of 1,109 observations from 334 patients with type I GD, after excluding 111 patients with undocumented values for chitotriosidase activity and/or serum CCL18 concentration, and 18 patients with deficient chitotriosidase activity (Figure 1). The median number of observations per patient was 2 (range: 1-14).

IPD were not available for the other 14 eligible primary studies that included 11 academic observational studies8,9,14,37,42(201 patients) and three industry-sponsored clinical trials of ERT or substrate reduction therapy44-46 (Online Supplementary Appendix S5). Overall, these studies enrolled 565 patients with GD, although overlap in study populations could not be excluded for five studies carried out in the Netherlands8,9,14,37,42 (201 patients) and four studies performed in Spain35,36,47,48 (203 patients).

Study characteristics

Among the nine studies that supplied IPD, six were international industry-funded clinical trials of ERT28-31,34,35 and three were observational in design17,32,33 (Online Supplementary Appendix S4). The study enrollment periods extended from 2003 to 2015, and the length of the follow-up ranged from 12 to 132 months for longitudinal studies. The median number of primary study participants and primary outcomes assessed by independent blinded reviewers in five studies, and undocumented in one study. The median values in primary studies ranged from 0.8-1.7 MN for liver volume, 2.7-14.1 MN for spleen volume, 11.7-14.0 g/dL for hemoglobin concentration, and 82-260 x10⁷/L for platelet count (Online Supplementary Appendix S6).

Study quality assessment

Six of the eight studies that evaluated the primary composite outcome fulfilled five or more QUADAS-2 tool criteria (Online Supplementary Appendix S7). The other two studies met four QUADAS-2 tool criteria. As data were collected by retrospective chart review, the Yale’s National Gaucher Disease Treatment Center study33 was considered at high risk of bias for index tests and pre-specified surrogate outcome assessment. This study was also considered at high risk of flow bias due to undocumented chitotriosidase activity or serum CCL18 concentration in 65% (115 of 167) of participants. In a randomized controlled trial that enrolled only treatment-naive patients, the applicability concern was high and it was not possible to formally determine whether biomarker assessment was blinded to the pre-specified outcomes and which analytical methods were used.34

Chitotriosidase activity and serum CCL18 concentration according to the outcomes

In one-stage meta-analysis involving 492 observations nested within 177 participants from eight primary studies (Figure 1), the primary composite outcome was associated with increased geometric mean of chitotriosidase activity (7,623 vs. 1,478 nmol/mL/h; geometric mean ratio: 5.29, 95% CI: 4.24-6.61, P<0.001) and CCL18 concentration (679 vs. 198 ng/mL; geometric mean ratio: 3.50, 95% CI: 2.57-5.58, P<0.001) compared with the absence of outcome (Table 1). Despite substantial between-study heterogeneity, the two-stage meta-analysis yielded comparable results, using random-effect models (Online Supplementary Appendices S8-9).

The effect size was quite homogeneous among the individual components of the primary composite outcome, with geometric mean ratio point estimates ranging from 2.96-5.43 for chitotriosidase activity and from 2.28-3.22 for serum CCL18 concentration (Table 1). Similar results were obtained for the secondary composite outcome and its individual components (Table 1), except for serum CCL18 concentration, which did not differ according to severe anemia (445 vs. 666 ng/mL, geometric mean ratio, 1.39, 95% CI: 0.82-2.37, P=0.22).

The geometric means of chitotriosidase activity (3,556 vs. 1,618 nmol/mL/h; geometric mean ratio: 1.47, 95% CI: 0.78-2.79, P=0.24) and serum CCL18 concentration (786 vs. 449 ng/mL; geometric mean ratio: 1.22, 95% CI: 0.81-1.81, P=0.34) did not differ according to symptomatic
bone events with imaging confirmation, among 218 observations nested within 97 participants from a single primary study (Table 1). The decrease in CCL18 concentration was paralleled by a similar trend in chitotriosidase activity over the 24 months of follow-up, among participants enrolled in four industry-sponsored clinical trials evaluating enzyme replacement therapy (ERT) (Online Supplementary Appendix S10).

Comparative accuracy of chitotriosidase activity and CCL18 concentration for assessing GD severity

In one-stage meta-analysis, area under the curve receiver operating characteristics (AUC-ROC) curve point estimates were 0.82 for chitotriosidase activity and 0.84 for serum CCL18 concentration for discriminating patients with GD according to the primary composite outcome (summary difference: 0.02, 95% CI: -0.02 to 0.05, P=0.82, Table 2). Adding chitotriosidase activity did not improve serum CCL18 concentration accuracy, as indicated by the AUC-ROC curve estimates (0.85 with and 0.84 without chitotriosidase activity, P=0.18) (Figure 2).

The summary estimates of the difference in the AUC-ROC curves were consistent in the two-stage meta-analysis (Online Supplementary Appendix S14). No evidence of

 tentative associations between CCL18 concentration and other outcomes were found.

Table 1. One-stage unpaired comparisons of chitotriosidase activity and CCL18 concentration according to pre-specified outcomes among type I Gaucher disease patients.

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>No.</th>
<th>Geometric mean (95% CI)</th>
<th>Mean ratio (95% CI)*</th>
<th>P</th>
<th>Geometric mean (95% CI)</th>
<th>Mean ratio (95% CI)*</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary composite outcome†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No outcome</td>
<td>212</td>
<td>1,478 (1,235 - 1,768)</td>
<td>1.00 (…)</td>
<td>&lt;0.001</td>
<td>198 (177 - 221)</td>
<td>1.00 (…)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>≥ 1 outcome</td>
<td>280</td>
<td>7,623 (6,520 - 8,913)</td>
<td>5.29 (4.24 - 6.81)</td>
<td></td>
<td>679 (612 - 755)</td>
<td>3.04 (2.57 - 3.58)</td>
<td></td>
</tr>
<tr>
<td>Secondary composite outcome†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No outcome</td>
<td>391</td>
<td>2,701 (2,349 - 3,166)</td>
<td>1.00 (…)</td>
<td>&lt;0.001</td>
<td>311 (283 - 342)</td>
<td>1.00 (…)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>≥ 1 outcome</td>
<td>101</td>
<td>13,516 (10,143 - 18,011)</td>
<td>4.35 (3.35 - 5.64)</td>
<td></td>
<td>1,050 (879 - 1,254)</td>
<td>3.05 (2.53 - 3.68)</td>
<td></td>
</tr>
<tr>
<td>Hemoglobin concentration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 11 g/dL</td>
<td>934</td>
<td>3,136 (2,836 - 3,468)</td>
<td>1.00 (…)</td>
<td>&lt;0.001</td>
<td>406 (381 - 432)</td>
<td>1.00 (…)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>&lt; 11 g/dL</td>
<td>102</td>
<td>10,984 (7,841 - 15,386)</td>
<td>2.96 (2.44 - 3.59)</td>
<td></td>
<td>1,057 (883 - 1,265)</td>
<td>2.28 (1.97 - 2.64)</td>
<td></td>
</tr>
<tr>
<td>Hemoglobin concentration</td>
<td></td>
<td></td>
<td></td>
<td>0.22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 8 g/dL</td>
<td>1,029</td>
<td>3,509 (3,178 - 3,876)</td>
<td>1.00 (…)</td>
<td></td>
<td>445 (418 - 473)</td>
<td>1.00 (…)</td>
<td></td>
</tr>
<tr>
<td>&lt; 8 g/dL</td>
<td>7</td>
<td>17,520 (8,076 - 38,005)</td>
<td>2.02 (1.00 - 4.10)</td>
<td></td>
<td>666 (331 - 1,340)</td>
<td>1.39 (0.82 - 2.37)</td>
<td></td>
</tr>
<tr>
<td>Platelet count</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 100 x10⁹/μL</td>
<td>758</td>
<td>2,413 (2,156 - 2,700)</td>
<td>1.00 (…)</td>
<td>&lt;0.001</td>
<td>343 (321 - 368)</td>
<td>1.00 (…)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>&lt; 100 x10⁹/μL</td>
<td>313</td>
<td>8,495 (7,339 - 9,833)</td>
<td>4.03 (3.46 - 4.70)</td>
<td></td>
<td>856 (782 - 957)</td>
<td>2.71 (2.42 - 3.03)</td>
<td></td>
</tr>
<tr>
<td>Platelet count</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 50 x10⁹/μL</td>
<td>958</td>
<td>3,121 (2,825 - 3,449)</td>
<td>1.00 (…)</td>
<td>&lt;0.001</td>
<td>412 (387 - 438)</td>
<td>1.00 (…)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>&lt; 50 x10⁹/μL</td>
<td>113</td>
<td>8,890 (6,569 - 12,030)</td>
<td>2.94 (2.32 - 3.73)</td>
<td></td>
<td>928 (786 - 1,094)</td>
<td>2.15 (1.81 - 2.55)</td>
<td></td>
</tr>
<tr>
<td>Liver volume</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 1.25 MN</td>
<td>447</td>
<td>2,242 (1,954 - 2,572)</td>
<td>1.00 (…)</td>
<td>&lt;0.001</td>
<td>317 (291 - 346)</td>
<td>1.00 (…)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>≥ 1.25 MN</td>
<td>240</td>
<td>10,181 (8,724 - 11,880)</td>
<td>3.96 (3.30 - 4.76)</td>
<td></td>
<td>886 (800 - 980)</td>
<td>2.59 (2.26 - 2.96)</td>
<td></td>
</tr>
<tr>
<td>Liver volume</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 2.5 MN</td>
<td>678</td>
<td>3,700 (3,290 - 4,161)</td>
<td>1.00 (…)</td>
<td>0.001</td>
<td>447 (415 - 482)</td>
<td>1.00 (…)</td>
<td>0.002</td>
</tr>
<tr>
<td>≥ 2.5 MN</td>
<td>9</td>
<td>30,353 (16,362 - 56,308)</td>
<td>3.16 (1.59 - 6.27)</td>
<td></td>
<td>1,479 (1,019 - 2,145)</td>
<td>2.25 (1.35 - 3.74)</td>
<td></td>
</tr>
<tr>
<td>Spleen volume</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 5 MN</td>
<td>240</td>
<td>1,583 (1,338 - 1,873)</td>
<td>1.00 (…)</td>
<td>&lt;0.001</td>
<td>209 (188 - 223)</td>
<td>1.00 (…)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>≥ 5 MN</td>
<td>265</td>
<td>8,111 (6,933 - 9,488)</td>
<td>5.43 (4.45 - 6.63)</td>
<td></td>
<td>740 (666 - 822)</td>
<td>3.22 (2.77 - 3.75)</td>
<td></td>
</tr>
<tr>
<td>Spleen volume</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 15 MN</td>
<td>434</td>
<td>2,871 (2,515 - 3,276)</td>
<td>1.00 (…)</td>
<td>&lt;0.001</td>
<td>335 (306 - 367)</td>
<td>1.00 (…)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>≥ 15 MN</td>
<td>71</td>
<td>18,539 (13,323 - 25,796)</td>
<td>4.14 (3.08 - 5.56)</td>
<td></td>
<td>1,059 (879 - 1,265)</td>
<td>3.21 (2.59 - 3.97)</td>
<td></td>
</tr>
<tr>
<td>Symptomatic bone events†</td>
<td></td>
<td></td>
<td></td>
<td>0.24</td>
<td></td>
<td></td>
<td>0.34</td>
</tr>
<tr>
<td>No</td>
<td>206</td>
<td>1,618 (1,335 - 1,961)</td>
<td>1.00 (…)</td>
<td></td>
<td>449 (402 - 501)</td>
<td>1.00 (…)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>12</td>
<td>3,556 (1,894 - 6,674)</td>
<td>1.47 (0.78 - 2.79)</td>
<td></td>
<td>786 (407 - 1,518)</td>
<td>1.22 (0.81 - 1.81)</td>
<td></td>
</tr>
</tbody>
</table>

*Summary geometric mean ratios and P-values for unpaired comparisons were derived from 3-level random intercept regression models for continuous dependent variables, with observations nested within patients and studies. The primary outcome was a composite of hemoglobin concentration < 11 g/dL, < 10 g/dL, and liver volume > 5 MN; and liver volume > 1.25 MN. Patients with splenectomy were excluded from this analysis. The secondary outcome was a composite of chitotriosidase activity, ≤ 15 MN, and liver volume > 2.5 MN. Patients with splenectomy were excluded from this analysis. # Osteonecrosis or fracture with imaging confirmation within the previous 12 months.
selective reporting for the primary composite outcome was found graphically (Online Supplementary Appendix 12) or statistically (P=0.20). However, this result should be interpreted with caution, given the limited number of primary studies.

Comparable effect sizes were observed for the secondary composite outcome and most of the individual components (Table 2). However, the AUC-ROC curve of serum CCL18 concentration for predicting severe anemia was lower than that of chitotriosidase activity (0.64 vs. 0.81, summary difference, -0.17, 95% CI: -0.33 to -0.08, P=0.008).

Subgroup analysis

No evidence of heterogeneity in the biomarker levels associated with the primary composite outcome was observed across the age groups and provision of ERT (Online Supplementary Appendix S13). Conversely, lower-quality studies (<5 QUADAS-2 criteria) displayed higher geometric mean ratios for chitotriosidase activity, according to the primary outcome (Online Supplementary Appendix S13).

Differences in the AUC-ROC curves for the secondary composite outcome were homogeneous across the age groups, QUADAS-2 criteria, fluorogenic substrate, and CCL18 assay type (Online Supplementary Appendix S14). However, serum CCL18 concentration was more accurate than chitotriosidase activity in discriminating patients with the primary composite outcome among those receiving ERT (Online Supplementary Appendix S14). No evidence of heterogeneity in serum CCL18 concentration accuracy as a function of chitotriosidase activity deficiency was found (Online Supplementary Appendices S15-16).

Sensitivity analysis

In leave-one-out sensitivity analysis that iteratively removed one study at a time, geometric mean ratios (Online Supplementary Appendix 17) and differences in the AUC-ROC curves (Online Supplementary Appendix 18) for the two biomarkers remained unchanged relative to the primary outcome, confirming that our results were not driven by any single study. In the additional sensitivity analysis, coding splenectomy as splenomegaly did not modify the associations of the primary and secondary composite outcomes with the two biomarkers (Online Supplementary Appendix S19), although the AUC-ROC curve for chitotriosidase activity decreased significantly (Online Supplementary Appendix S20).

Discussion

Most published studies reporting on chitotriosidase activity and CCL18 concentration accuracy in assessing GD severity are of relatively limited sample size. In this context, our meta-analysis summarizes evidence from 1,109 observations nested within 334 participants enrolled in nine primary studies, with a broad range of patient age and disease severity.

Our analysis indicates that CCL18 concentration is as accurate as chitotriosidase activity in discriminating patients on the basis of our primary composite outcome. This finding was robust in the sensitivity analysis, and we did not find evidence of between-study heterogeneity (P=0%, P=0.62) in two-stage meta-analysis. Additionally, we did not detect any significant heterogeneity among studies and patient subgroups, although CCL18 concentration may be more accurate in patients receiving ERT. Our primary result is also supported by consistency in the effect sizes for most of the individual outcome components and the secondary composite outcomes.

Our meta-analysis corroborates and extends evidence from previous studies showing that chitotriosidase activity and CCL18 concentration relate to visceral and hematological parameter abnormalities in treated and untreated patients.5-12,15 A noteworthy exception was severe anemia (<8 g/dL), for which CCL18 showed a lower AUC-ROC curve than chitotriosidase activity. This observation is consistent with our failure to show a significant difference in the CCL18 level for patients with and without severe anemia. Interestingly, weak or variable associations were previously reported between hemoglobin concentration and chitotriosidase activity or CCL18 concentration.12-15,16 A potential explanation might relate to the failure of physicians to correctly diagnose concurrent causes of severe anemia in GD patients. Yet, this cannot explain why the accuracy differed between chitotriosidase activity and CCL18 concentration. Due to the limited number of severe anemia cases in our meta-analytical sample (i.e., 7 among the 1,036 observations from 309 participants), our findings should be interpreted with caution and deserve confirmation in independent samples. Consistent with previous studies,5,10 we found that splenectomy altered the accuracy of chitotriosidase activity in discriminating patients with the primary outcome.

Another striking result of our meta-analysis is that the addition of chitotriosidase activity did not improve the
accuracy of serum CCL18 concentration in identifying patients with the primary outcome, as shown by the unchanged AUC-ROC curve. This negative finding suggests that combining chitotriosidase activity and serum CCL18 concentration would have a limited incremental value, probably because they convey redundant information on Gaucher cell burden.14,40-51 The recent glucosylsphingosine biomarker9 may provide complementary information about the pathophysiological process in GD.14,38 Indeed, plasma glucosylsphingosine relates to sphingolipid turnover while chitotriosidase activity and CCL18 concentration are indicative of overall Gaucher cell mass. Glucosylsphingosine appeared to be highly sensitive and specific for the primary diagnosis and severity assessment of GD in previous studies.35-37 It would be interesting to compare the accuracy of glucosylsphingosine and CCL18 concentration using the same approach as in our meta-analysis although insufficient data exist to perform this analysis to date. Meanwhile a prospective comparative study of the three biomarkers is warranted to evaluate their contribution in monitoring GD activity.

In contrast to CCL18, glucosylsphingosine can be measured in dried blood spot and has emerged as a second-tier biomarker for newborn screening for GD.51 Measuring biomarkers in dried blood spot samples might facilitate GD activity monitoring for clinics where the technology is not available. A limitation of our primary composite outcome is that it did not capture bone manifestations, despite their major impact on the quality of life.2,24 Previous studies reported that chitotriosidase activity level correlates with bone complications,15 history of osteonecrosis,17 and number of anatomical sites of osteonecrosis,17 although these associations were inconsistent.15,23 Similarly, CCL18 concentration has been inconsistently associated with history of osteonecrosis and number of anatomical sites of osteonecrosis.13,17 Unfortunately, information on symptomatic bone events was available in only one study included in our meta-analysis,13 and we could not reach a definitive conclusion for this outcome. There are several potential explanations for the seemingly inconsistent findings between our analysis of bone events and the study by Deegan et al.17 First, we used a stricter definition of bone manifestations, which included symptomatic major events (i.e., skeletal fracture, osteonecrosis, or avascular necrosis) that could be dated and excluded progressive alterations (i.e., osteoporosis, Erlenmeyer flask femur deformity). Second, we took into account the data temporality by including bone events that occurred within the previous 12 months of biomarker analysis while Deegan et al.15 analyzed the previous history of bone events that occurred during the lifetime. Third, the unit of analysis was different between our analysis (218 observations nested within 97 patients) and the study by Deegan et al.17 (100 patients).

Two other clinical trials included in our meta-analysis recorded bone mineral density,52,53 but they did not report clinically relevant endpoints for bone disease (i.e., skeletal fracture, osteonecrosis, or avascular necrosis). Thus, the imprecise definition of diagnostic terms, the lack of standardized criteria for bone assessment, and the slow improvement in bone outcomes after specific treatment probably hamper the identification of a relationship between biomarkers and this important clinical outcome.34,55 Another approach in assessing skeletal disease is to measure residual biomarker levels after remission of visceral disease (reversal of hepatosplenomegaly in splenectomized patients or reversal of hepatosplenomegaly in intact spleen patients).

Our findings have clinical implications for GD monitoring. Due to the comparable accuracy of the two biomarkers, the impact of common genetic polymorphism on chitotriosidase activity and the analytical robustness of CCL18 assays, it should be logical to recommend measuring CCL18 concentration rather than chitotriosidase activity as a biomarker of Gaucher cell burden and visceral and hematological damage. Measuring chitotriosidase activity in addition to CCL18 concentration does not appear useful in the routine practice or in clinical research. Indeed, the two biomarkers are strongly correlated and convey redundant information on

<table>
<thead>
<tr>
<th>Table 2. One-stage paired comparisons of AUC-ROC curves for chitotriosidase activity and CCL18 concentration for discriminating type I Gaucher disease patients with prespecified outcomes.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome</td>
</tr>
<tr>
<td>-------------------------------</td>
</tr>
<tr>
<td>Primary composite outcome†</td>
</tr>
<tr>
<td>Secondary composite outcome†</td>
</tr>
<tr>
<td>Hemoglobin concentration &lt; 11 g/dL</td>
</tr>
<tr>
<td>Hemoglobin concentration &gt; 11 g/dL</td>
</tr>
<tr>
<td>Platelet count &lt; 100x10⁹/L</td>
</tr>
<tr>
<td>Platelet count &gt; 50x10⁹/L</td>
</tr>
<tr>
<td>Liver volume &gt; 1.25 MN</td>
</tr>
<tr>
<td>Liver volume &gt; 2.5 MN</td>
</tr>
<tr>
<td>Spleen volume &gt; 5 MN</td>
</tr>
<tr>
<td>Spleen volume &gt; 15 MN</td>
</tr>
<tr>
<td>Symptomatic bone events†</td>
</tr>
</tbody>
</table>

AUC-ROC: area under the curve receiver operating characteristics; CI: Confidence Interval; MN: multiple of normal. Summary estimates for area under ROC curves and P-values for paired comparisons were derived from non-parametric ROC analysis with bootstrap resampling that accounted for observation clustering within patients and primary studies. The primary outcome was a composite of hemoglobin concentration < 11 g/dL (< 10 g/dL for patients 12 to 59 months of age), platelet count < 100x10⁹/L, spleen volume > 5 MN, and liver volume > 1.25 MN. Patients with splenectomy were excluded from this analysis. The secondary outcome was a composite of hemoglobin concentration < 8 g/dL (< 7 g/dL for patients 12 to 59 months of age), platelet count < 50x10⁹/L, spleen > 15 MN, and liver volume > 2.5 MN. Patients with splenectomy were excluded from this analysis. *Summary estimates for area under ROC curves and P-values for paired comparisons were derived from non-parametric ROC analysis with bootstrap resampling that accounted for observation clustering within patients and primary studies. †The primary outcome was a composite of hemoglobin concentration < 11 g/dL (< 10 g/dL for patients 12 to 59 months of age), platelet count < 100x10⁹/L, spleen volume > 5 MN, and liver volume > 1.25 MN. Patients with splenectomy were excluded from this analysis. The secondary outcome was a composite of hemoglobin concentration < 8 g/dL (< 7 g/dL for patients 12 to 59 months of age), platelet count < 50x10⁹/L, spleen > 15 MN, and liver volume > 2.5 MN. Patients with splenectomy were excluded from this analysis. **Summary estimates for area under ROC curves and P-values for paired comparisons were derived from non-parametric ROC analysis with bootstrap resampling that accounted for observation clustering within patients and primary studies.
GD severity. Additionally, chitotriosidase activity has limited incremental value in improving CCL18 concentration accuracy, as reflected by the unchanged AUC-ROC curve estimate.

Our meta-analysis demonstrates that chitotriosidase activity and CCL18 concentration relate to visceral and hematological parameters that can be easily monitored in the routine practice. Hence, additional studies are needed to further investigate the usefulness of these two biomarkers. Lastly, the relationships between GD biomarkers and patient-centered outcomes, including quality of life, fatigue, chronic pain, and restriction of daily activities should be evaluated.

This meta-analysis analyzed IPD from nine primary studies that were completed since 2004, summarizing the most recent available evidence on the accuracy of chitotriosidase activity and CCL18 concentration in assessing type 1 GD severity. All studies included in this meta-analysis evaluated both biomarkers in the same patients, providing unconfounded comparative accuracy estimates. Moreover, our findings have strong generalizability because we combined primary studies that enrolled various populations of patients worldwide.

However, our meta-analysis also has a few caveats that must be considered. First, the finding interpretation is inevitably limited by the unavailability of IPD from the other 14 potentially eligible primary studies. Second, there was substantial between-study heterogeneity in chitotriosidase activity and, to a lesser extent, in CCL18 concentration. The lack of assay standardization may have contributed to this heterogeneity, although chitotriosidase activity and CCL18 concentration were measured at a single central core laboratory in five clinical trials. Therefore, we used mixed-effect regressions to account for this heterogeneity and performed sensitivity and subgroup analyses that supported the robustness of the summary estimates.

In conclusion, CCL18 concentration is as accurate as chitotriosidase activity in assessing hematological and visceral parameters of GD activity and can be measured in patients who are homozygous for null chitotriosidase variants. However, the observed between-study heterogeneity and the limitations of this meta-analysis should encourage the international community to i) implement a strategy of homogenization and standardization of dosage techniques, more than 15 years after their implementation in clinical practice and ii) set up a prospective large-scale study to evaluate GD biomarkers, including glucosyl-sphingosine. This latter biomarker seems interesting from a pathophysiological perspective, but its superiority over chitotriosidase activity and CCL18 concentration remains to be documented.

Disclosures
PD received speaker and board membership fees from Takeda and consulting fees from Sanofi Corporation; PM received research grant, lecture fee honoraria, and travel support from Sanofi-Genzyme; AZ received consulting fees from Shire and Prevail Therapeutics, honoraria from Shire Corporation and Pfizer, he is an employee at the Gaucher Clinic, Shaare Zedek Medical Center, Israel, which received research grants from Shire Corporation, Centogene and Pfizer, and support from Shire Corporation, Pfizer and Sanofi-Genzyme for participation in their respective registries (GOS, TALIAS and ICGG). MGB received speaker fees and unrestricted research grants from Sanofi-Genzyme and Shire Corporation. The other authors have no conflict of interest relevant to this study.

Contributions
TR contributed to the study design, data acquisition, interpretation of the results, and manuscript preparation; PD, EP, PM, RG, AZ, JB, CB contributed to data acquisition and manuscript preparation; BP and JL contributed to the study design, data acquisition, data management, statistical analysis, interpretation of the results, and manuscript preparation; MB provided project leadership, contributed to the interpretation of the results, and manuscript preparation. All authors approved the final version of the manuscript.

Acknowledgments
The authors are indebted to the investigators of all primary studies for their contributions that allowed carrying out this secondary analysis. They are also grateful to Zoya Panahloo, Farid Merati, Hak-Myung Lee, Bjorn Mellgard, Michael Craig, and Hartmann Wellhoefer from Shire, Lexington, MA, for sharing individual participant data from clinical trials and facilitating the pooled analysis. Statistical analysis was performed within the framework of the Grenoble Alpes Data Institute (ANR-15-IDEX-03).

References
Biomarkers in Gaucher disease


