

# Factors Contributing to the Efficacy-Effectiveness Gap in the Case of Orphan Drugs for Metabolic Diseases

Y. Schuller<sup>1</sup> · C. E. M. Hollak<sup>1</sup> · C. C. Gispen-de Wied<sup>2</sup> · V. Stoyanova-Beninska<sup>2</sup> · M. Biegstraaten<sup>1</sup>

Published online: 27 July 2017

© The Author(s) 2017. This article is an open access publication

## Abstract

**Introduction** Authorization of orphan medicinal products (OMPs) is often based on studies with several methodological shortcomings. Hence, data are difficult to interpret and efficacy does not always correspond to real-world effectiveness. We investigated to what extent an efficacy-effectiveness gap exists for OMPs for metabolic diseases and set out to explore which factors contribute to it.

**Methods** We included all OMPs for rare metabolic diseases authorized in the EU up to 1 January 2016. Efficacy data were obtained from European Public Assessment Reports, relative effectiveness data from the Dutch National Healthcare Institute website, and real-world effectiveness data from literature and interviews with experts and patients. Efficacy and effectiveness were scored as ‘no effect’, ‘unclear’ or ‘good’ based upon a prespecified scoring system.

**Results** We identified 31 authorized OMPs, of which 21 had post-marketing studies available, thus making it possible to score real-world effectiveness. Eight of 21 (38%) OMPs had a ‘good’ real-world effectiveness. The use of a clinical or validated surrogate primary endpoint and a

representative study population seemed to be related to good effectiveness in the real world, as were type of marketing authorization, study population and disease prevalence.

**Conclusions** This study revealed that less than half of the authorized OMPs are effective in the real world. Since the type of primary endpoint used in the pivotal study seems to be associated with good real-world effectiveness, it is important to agree upon study endpoints through early dialogues among relevant stakeholders.

## Key Points

Less than half of the authorized orphan drugs for metabolic diseases show good effectiveness in the real world.

Of drugs with an unclear efficacy at the time of authorization only 21% had good real-world effectiveness.

The use of a clinical or validated surrogate primary endpoint in the pivotal study seems to be the most important factor associated with good real-world effectiveness.

**Electronic supplementary material** The online version of this article (doi:10.1007/s40265-017-0788-z) contains supplementary material, which is available to authorized users.

✉ Y. Schuller  
y.schuller@amc.nl

<sup>1</sup> Department of Endocrinology and Metabolism, Academic Medical Center, University of Amsterdam, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands

<sup>2</sup> College ter Beoordeling Geneesmiddelen/Medicines Evaluation Board, Graadt van Roggenweg 500, 3531 AH Utrecht, The Netherlands

## 1 Introduction

In the European Union (EU), a disease is considered orphan if it is a life-threatening or seriously debilitating disorder, with a prevalence of less than 5 in 10,000. Up to

now, about 7000 rare diseases have been documented, affecting 30–40 million people in the EU [1]. In the past 16 years, 130 orphan drugs have been marketed. Initially, the small consumer's market offered little attraction to the pharmaceutical industry to develop drugs for rare diseases. To stimulate the development of orphan drugs, European legislation was introduced in 2000. The legislation implies that, for a product with orphan status, the pharmaceutical industry is entitled to receive (1) a fee reduction for scientific advice, (2) reduction of fees for marketing authorization application, (3) centralized registration in the EU, and (4) 10 years' market exclusivity after marketing authorization [2]. Since the implementation of the Orphan Drug Legislation and its incentives, orphan medicinal products (OMPs) are being developed at an increasing rate. In 2015, for example, the European Medicines Agency (EMA) approved 14 OMPs, whereas in 2000 only three OMPs were approved [3]. In addition, more than 40% of all registered OMPs have been developed in the past 5 years [4]. Expectations are that the development rate of OMPs will only increase in the future [5].

Since patients with rare diseases have the same right to access to effective and safe drugs as all other patients, robust evidence is needed regarding benefits and risks. At evaluation of OMPs, however, regulatory authorities are often confronted with difficulties that are inherent to the rarity of the diseases for which the drugs are developed. Interpretation of data is often hampered, because clinical trials may lack a placebo control, are open label or are performed in a limited number of patients. Moreover, the heterogeneity of each disease and the frequent use of non-validated biomarkers or functional outcome measures (FOMs) may hamper reliable interpretation of data [6–8]. It is important to emphasize that biomarkers or FOMs are only 'surrogate endpoints' if they are validated (Box 1).

The regulatory authorities assess the benefit/risk ratio of OMPs based on the evidence provided in the dossier submitted to the EMA. In case the evidence is insufficient to conclude on a positive benefit/risk ratio and thereby to grant a full marketing authorization, the EU legislation allows two alternatives. The first is a 'conditional approval', which can be granted for drugs when the clinical data are limited, but indicate a positive benefit/risk ratio, and when it is likely that additional data will be provided in a reasonable timeframe. If the additional data confirm the positive benefit/risk ratio, the conditional approval can be changed to a full marketing authorization. The second is a marketing authorization under 'exceptional circumstances', which is granted in cases when limited data are provided at the time of marketing authorization application, but additional data cannot reasonably be expected to be delivered in the future due to for example the rarity of the disease [9]. Nonetheless, the marketing authorization holder (MAH) is

generally requested by the regulatory authorities to collect additional data on safety and long-term effectiveness in the post-marketing period.

After marketing authorization, regulatory authorities in each EU member state make their own assessment before granting reimbursement. These national assessments may for example take into account relative effectiveness—the 'net benefit' of the new drug needs to be desirable, sufficiently big and a relevant addition to the present care—probability of real-world effectiveness as well as cost-effectiveness. The debate on real-world effectiveness has recently been fueled by the authorization of a drug for mucopolysaccharidosis (MPS) IV, elosulfase alfa (Vimizim®). It was granted a full marketing authorization by the EMA on the basis of a favorable benefit/risk ratio regarding quality, safety and efficacy data submitted by the manufacturer [10]. This was mainly based on a short-term study that showed a significant increase in mobility, using the 6-min walk test (6-MWT). However, it was rejected for reimbursement by several national authorities (e.g., The Netherlands, Sweden, Belgium and Spain). The Dutch reimbursement authority (Zorginstituut Nederland, ZiN) stated that the data are limited to a suboptimal short-term outcome in a heterogeneous population, showing a small effect with a wide confidence interval that is not clinically relevant, herewith estimating a small probability of good real-world effectiveness.

This difference between the opinion of the EMA (or other regulatory authorities such as the US Food and Drug Administration) and the (estimated) real-world effectiveness is in the literature often referred to as the 'efficacy-effectiveness gap' and is particularly an issue in the field of orphan drugs [11]. To support the harmonized decisions of the EMA and the national reimbursement authorities in the future, it is important to bridge this gap. Hence, the purpose of our research project is to investigate whether data used for authorization (i.e. data from the pivotal studies) are predictive for real-world effectiveness (i.e. effect on clinical endpoints in clinical practice and post-marketing), and to explore which factors contribute to the efficacy-effectiveness gap. In this article we focus on OMPs for metabolic diseases. Based on the results of this study we aim to provide recommendations to be used by both investigators and regulators in order to reduce this gap and improve both access to and appropriate use of such medicines.

## 2 Methods

The 'Community register of orphan medicinal products for human use' of the European Commission was used to identify all OMPs that were authorized for the treatment of metabolic diseases up to 1 January 2016 [14]. A detailed

**Box 1: Definitions of endpoints**

Several terms are used to describe measurements of disease and treatment effects, such as biomarkers, surrogate markers, surrogate endpoints, intermediate endpoints, etc. We applied the definitions of the Biomarker Definitions Working group [12]. The definition of ‘functional outcome measure’ was inspired by an article on Duchenne Muscular Dystrophy [13]

*Biomarker* A characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes or pharmacological responses to a therapeutic intervention

*Functional outcome measure* A characteristic that is measured by either a function test or a (disease) symptom score

*Surrogate endpoint* A biomarker or functional outcome measure that is intended to substitute for a clinical endpoint and is expected to predict clinical benefit (or harm or lack of benefit or harm). Only a small subset of biomarkers and functional outcome measures may achieve surrogate endpoint status based on epidemiological, therapeutic, pathophysiological or other scientific evidence

*Clinical endpoint* A characteristic or variable that reflects how a patient feels, functions or survives

description of study methods can be found in Supplementary Table 1.

**2.1 Pre-Marketing Data**

To collect information on the pivotal studies (pre-marketing data), European Public Assessment Reports (EPARs) were used, which are available on the website of the EMA. We assessed the quality of evidence of the pivotal studies using the COMPASS (Clinical evidence of Orphan Medicinal Products—an ASSESSment) tool [15]. In addition, we scored the efficacy of the OMPs based on the effect on primary and secondary endpoints in the pivotal studies and the considerations of the EMA as set out in the EPARs as follows: category 3 = good efficacy, category 1 or 2 = unclear efficacy, and category 0 = no effect (Table 1).

**2.2 Post-Marketing Data**

We divided the evaluation of effectiveness in two phases: (1) relative effectiveness at the time of reimbursement decisions by the Dutch National Health Care Institute, and (2) real-world effectiveness, including all post-marketing literature that was available until June 2016, supplemented with experts’ and patients’ opinions. If an OMP was authorized after January 2014 it was assumed that no reliable estimate of the real-world effectiveness could be made. Similar to pre-marketing data, the effectiveness of each OMP was divided into three categories as defined by the authors: category 3 = good effectiveness, category 2 or 1 = unclear effectiveness, or category 0 = no effect (Table 1). In addition, experts were interviewed and their opinions were used to up- or downgrade the category a drug received by one category at most. Also, expert opinions were used to select relevant post-marketing studies. Patients were approached to fill in a short online survey with questions about their main symptoms and their opinion on the choice of endpoints in clinical trials. Patients’ opinions were taken into consideration upon identifying relevant clinical endpoints.

**2.3 Statistical Analysis**

To explore which factors contribute to the efficacy-effectiveness gap, the variables of the COMPASS tool as well as the relative and real-world effectiveness were dichotomized and compared with  $2 \times 2$  tables. COMPASS variables were dichotomized as follows: type of primary endpoint was divided into clinical endpoints and surrogate endpoints versus non-validated biomarkers and FOMs, study population into extreme selection of patient population versus representative selection of patient population, type of marketing authorization into exceptional or conditional versus full authorization, study phase into phase III versus other phases, randomization into yes or no, and disease prevalence into rare or ultra-rare. Relative and real-world effectiveness were divided into ‘no or unclear effect’ (category 0, 1 or 2) or ‘good effect’ (category 3). Fisher’s exact test was used to assess statistical significance. A similar analysis was performed to compare efficacy [dichotomized into ‘no or unclear efficacy’ (category 0, 1 or 2) or ‘good efficacy’ (category 3)] and effectiveness. All analyses were performed using IBM SPSS Statistics version 22.

**3 Results****3.1 Orphan Medicinal Products (OMPs) and Indications**

We included 27 OMPs in our study, which were authorized for 25 metabolic orphan diseases (Table 2). Four OMPs were authorized for two different indications (miglustat, carglumic acid and pasireotide) or for different age groups (alglucosidase alfa), adding up to a total of 31 OMPs.

**3.2 Pre-Marketing Data—Characteristics of Pivotal Studies**

From the EPARs of the abovementioned 31 OMPs, 40 ‘pivotal’ or ‘main’ studies were identified (Table 3). Nineteen of these 40 (47%) studies were phase III studies,

**Table 1** Categories of efficacy and effectiveness

	Conclusion		Category
	Efficacy	Effectiveness	
Clinically relevant <sup>a</sup> effect on $\geq 1$ clinical or surrogate endpoint when compared with standard of care	Efficacy is good	Effectiveness is good	3
OMP is implemented in (inter)national state of the art guideline and quality of evidence is good	n/a	Effectiveness is good	3
Statistically significant effect on $\geq 1$ clinical endpoint or surrogate endpoint	Efficacy is unclear	Effectiveness is unclear	2
No study on effect on clinical endpoints, but a statistically significant effect on (non-validated) biomarker or functional outcome measure	Efficacy is unclear	Effectiveness is unclear	1
Studies provided by manufacturer fail to show clinically relevant or statistically significant effect on clinical endpoint(s)	No clinical effect	n/a	0
Post-marketing studies show that there is no clinically relevant or statistically significant effect on clinical endpoint(s)	n/a	No clinical effect	0
No reimbursement decision is made yet. No post-marketing studies (or authorization after January 2014)	n/a	Cannot be evaluated	n/a

<sup>a</sup> Defined as: The extent to which a new therapy contributes to a meaningful (positive) change in daily-life activities

n/a Not applicable

and the majority were multicentre (30/40, 75%) and multinational (25/40, 63%). In 26/40 (65%) studies the allocation was randomized and a control arm was used in 27/40 (67%). Nineteen of the 40 (47%) studies applied some type of blinding. In 6/40 (15%) studies, the study population was an extreme selection of the patient population due to strict inclusion criteria (i.e., only men included, considerable proportion of patients were physically unable to perform primary outcome measure test, etc.). Five of the 31 (16%) OMPs were authorized on the basis of an improvement on a clinical endpoint, while changes in FOMs or biomarkers resulted in the approval of 6/31 (19%) and 20/31 (65%) OMPs, respectively. Of the latter 20, seven used biomarkers that were classified as surrogate endpoints.

### 3.3 Pre-Marketing Data–Efficacy

Category 3 (good efficacy, see Table 1) was assigned to 11/31 (35%) OMPs, category 2 (unclear efficacy) to 9/31 (30%) and category 1 (unclear efficacy) to 11/31 (35%) (Table 2). Supplementary Table 2 gives an overview of the EMA considerations upon authorizing the individual OMPs, including the post-marketing commitments as agreed upon with the MAHs.

### 3.4 Post-Marketing Data–Relative Effectiveness at the Time of Reimbursement Decisions

Eleven OMPs were not yet assessed by the Dutch National Health Care Institute for one of the following reasons: (1) marketing authorization was granted only recently, (2) the

OMP did not carry a claim of added benefit, or (3) the OMP did not have an annual budget impact exceeding 2.5 million euros [5]. Of the remaining 20 OMPs, 7/20 (35%) were classified as category 3 (good effectiveness), while 13/20 (65%) OMPs were categorized as category 1 or 2 (unclear effectiveness) (Table 2).

### 3.5 Post-Marketing Data–Real-World Effectiveness

For ten OMPs either no post-marketing studies were performed or a reliable judgement about effectiveness could not be made since marketing authorization was granted after January 2014 (Table 2). Of the 21 OMPs with post-marketing studies available, 8/21 (38%) were classified as category 3 (good effectiveness), 8/21 (38%) as category 2 (unclear effectiveness) and 4/21 (19%) as category 1 (unclear effectiveness). One OMP (5%) was categorized as category 0 (no effect). See Supplementary Table 3 for an overview of post-marketing evidence used. Forty-one experts participated in an in-depth interview, resulting in a range of 0–3 experts per OMP. For two OMPs no experts were willing to participate. Eight experts were involved in the pivotal trials of eight OMPs. Of these, three OMPs were not included in our analysis since post-marketing data were not (yet) available. Following expert opinions, the judgement on real-world effectiveness was upgraded for two OMPs and downgraded for one OMP (Table 2). A sensitivity analysis without up-/downgrading the categories according to expert opinion did not change the results. Seven patient organizations (POs) were approached representing 25 orphan diseases and 27 OMPs. Four out of seven POs (57%) were willing to participate. The

**Table 2** Overview of orphan medicinal products (OMPs) for metabolic diseases authorized in the EU between 2000 and 2016

Generic name	Drug	Disease	Year of authorization	Type of authorization	Primary endpoint <sup>g</sup> in pivotal study	Category at time of authorization (EMA)	Category at time of reimbursement decision (ZIN)	Category real-world effectiveness
Miglustat	Zavesca <sup>a</sup>	1. Gaucher disease 2. Niemann-Pick C	2002 2006	Exceptional Exceptional	<b>Spleen and liver size</b> <i>HSEM velocity</i>	3 1	2 2	2 0 <sup>d</sup>
Velaglucerase alfa	Vpriv	Gaucher disease	2010	Full	<b>Hemoglobin concentration</b>	2	2	3 <sup>e</sup>
Eliglustat	Cerdelga	Gaucher disease	2015	Full	<b>Spleen volume, hematological parameters</b>	2	n/a	n/a
Agalsidase beta	Fabrazyme <sup>a</sup>	Fabry disease	2001	Exceptional <sup>f</sup>	GL-3 reduction in a kidney biopsy	1	3	2
Agalsidase alfa	Replagal <sup>a</sup>	Fabry disease	2001	Exceptional <sup>f</sup>	<b>Neuropathic pain, myocardial GL-3 levels</b>	3	2	2
Alglucosidase alfa	Myozyme	Pompe disease	2006	Full	Infants: <b>survival (2 studies)</b> Adults: <i>6-MWT and FVC</i>	3 1	3 1	3 2
Laronidase	Aldurazyme <sup>a</sup>	MPS 1 (Hunter disease)	2003	Exceptional <sup>f</sup>	<i>6-MWT and FVC</i>	1	1	2
Idursulfase	Elaprase	MPS 2 (Hunter disease)	2007	Exceptional	<i>6-MWT and FVC</i>	2	1	1
Elosulfase alfa	Vimizim	MPS 4a (Morquio disease)	2014	Full	<i>6-MWT</i>	1	1	n/a
Galsulfase	Naglazyme <sup>a</sup>	MPS 6 (Maroteaux-Lamy Syndrome)	2006	Exceptional	<i>12-MWT</i>	1	1	1
Sebelipase alfa	Kanuma	Lysosomal acid lipase deficiency	2015	Full	<b>Survival, ALAT normalization</b>	3	n/a	n/a
Sapropterin	Kuvan	Phenylketonuria	2008	Full	<b>Phenylalanine concentration in blood</b>	3	3	3
Betaine	Cystadane	Homocystinuria	2007	Full	<b>Homocysteine levels</b>	3	2	n/a
Mercaptamine	Procytsbi <sup>b</sup>	Cystinosis	2013	Full	<b>White blood cell cystine levels</b>	2	n/a	3
Nitisinone	Orfadin <sup>a</sup>	Hereditary tyrosinemia type 1	2005	Exceptional <sup>f</sup>	<b>Survival</b>	3	3	3
Alipogene tiparvoec	Glybera	Lipoprotein lipase deficiency	2012	Exceptional	Fasting triglyceride levels	3	n/a	n/a
Carglumic acid	Carbaglut <sup>a</sup>	1. Hyperammonemia (due to NAGS deficiency)	2003	Exceptional <sup>f</sup>	Plasma ammonia levels	3	3	3 <sup>f</sup>

Table 2 continued

Generic name	Drug	Disease	Year of authorization	Type of authorization	Primary endpoint <sup>g</sup> in pivotal study	Category at time of authorization (EMA)	Category at time of reimbursement decision (ZiN)	Category real-world effectiveness
		2. Hyperammonemia (due to organic acidurias)	2011	Full	Plasma ammonia levels	1	n/a	2
Zinc	Wilzin <sup>a</sup>	Wilson's disease	2004	Full	Copper levels	3	3	3
Cholic acid	Kolbam	Inborn errors in primary bile acid synthesis	2014	Exceptional	Bile acid levels and liver function	2	n/a	n/a
Cholic acid	Orphacol	Inborn errors in primary bile acid synthesis	2013	Exceptional	Bile acid levels and liver function	3	n/a	n/a
Hydrocortisone	Plenadren <sup>b</sup>	Adrenal insufficiency	2011	Full	Serum cortisol	1	n/a	1
Pasireotide	Signifor	1. Cushing's disease	2012	Full	Urinary free cortisol levels	1	1	1
		2. Acromegaly	2012	Full	Biochemical control (GH and IGF-I)	1	2	2
Ketoconazole	Ketoconazole	Cushing's syndrome	2015	Full	Plasma cortisol levels, urinary FC 24 h	1	3	3
Pegvisomant	Somavert <sup>a</sup>	Acromegaly	2002	Full	IGF-I concentrations	2	2	2
Glycerol phenylbutyrate	Ravicti	Urea cycle disorders	2015	Full	<b>Blood ammonia levels</b>	2	n/a	n/a
Asfotase alfa	Strensiq	Hypophosphatasia	2015	Exceptional	Rickets severity (2 studies)	2	n/a	n/a
Afamelanotide	Scenesse	Erythropoietic protoporphyria	2014	Exceptional	<b>Duration of sunlight exposure</b>	2	n/a	n/a

6-MWT 6-min walk test, 12-MWT 12-min walk test, ALAT alanine aminotransferase, EMA European Medicines Agency, FC free cortisol, FVC forced vital capacity, GH growth hormone, GL-3 globotriaosylceramide, HSEM horizontal saccadic eye movement, IGF-I insulin-like growth factor 1, n/a Not applicable, NAGS N-acetylglutamate synthetase, ZiN Zorginstituut Nederland (Dutch National Health Care Institute)

<sup>a</sup> Withdrawn from the community register of orphan medicinal products at the end of the 10-year period of market exclusivity

<sup>b</sup> 'Hybrid medicines', indicating that they are similar to a 'reference medicine' containing the same active substance (cysteamine and hydrocortisone). The pharmaceutical forms of both drugs have been used for many years, but the novel tablets are available in a formulation that allows for a delayed release of the active substance. Hereby, dosing frequency is reduced and adherence to treatment is improved

<sup>c</sup> Although the type of approval is classified as 'exceptional' in the EPAR, it changed to 'full approval' after additional data was provided by the marketing authorization holder, herewith implying that the original approval was actually 'conditional' instead of 'exceptional'

<sup>d</sup> Score was downgraded following expert opinion. However, original score was 0, hence final score did not change

<sup>e</sup> Score was upgraded following expert opinion. However, original score was 3, hence final score did not change

<sup>f</sup> Score was upgraded (from 2 to 3) following (impartial) expert opinion

<sup>g</sup> **Clinical endpoint:** (validated) surrogate; functional outcome measure; biomarker

**Table 3** Characteristics of pivotal studies (*n* = 40)

Drug	Study	Study phase	Multicentre	Multinational	Endpoint	Control	Randomized	Blinding	Study population representative	Duration	No. of patients	QoL-end point
Miglustat (Gaucher)	OGT 918-001	I/II	Yes	Yes	Surrogate <sup>a</sup>	No	No	n/a	Yes	12 months	28	No
Miglustat (NPC)	OGT 918-007	II	Yes	Yes	FOM	St. of care	Yes	No	Yes	12 months	31	Yes
Velaglucerase alfa	HGT-GCB-039	III	Yes	Yes	Surrogate <sup>a</sup>	Active comp.	Yes	Yes	Yes	9 months	35	No
Eliglustat	GZD02507	III	Yes	Yes	Surrogate <sup>a</sup>	Placebo	Yes	Yes	Yes	39 weeks	40	Yes
	GZGD02607	III	Yes	Yes	Surrogate <sup>a</sup>	Active comp.	Yes	No, but justified	Yes	52 weeks	160	No
Agalsidase beta	AGAL-1-002-98	III	Yes	Yes	Biomarker	Placebo	Yes	Yes	No	20 weeks	58	Yes
Agalsidase alfa	TKT-003	II	No	No	Clinical	Placebo	Yes	Yes	No	6 months	26	Yes
	TKT-005	II	No	No	Biomarker	Placebo	Yes	Yes	Yes	6 months	15	No
Alglucosidase alfa	AGLU01602	II/III	Yes	Yes	Clinical	Diff. dose	Yes	No	Yes	52 weeks	18	No
	AGLU01702	II	Yes	Yes	Clinical	Hist. control	No	n/a	Yes	120 weeks	21	No
	AGLU02704	NR	Yes	Yes	FOM	Placebo	Yes	Yes	Yes	78 weeks	90	Yes
Laronidase	ALID-003-99	III	Yes	Yes	FOM	Placebo	Yes	Yes	No	12 months	45	Yes
Idursulfase	TKT024	II/III	Yes	Yes	FOM	Placebo	Yes	Yes	No	12 months	96	No
Elosulfase alfa	MOR-004	III	Yes	Yes	FOM	Placebo	Yes	Yes	No	24 weeks	176	Yes
Galsulfase	ASB-03-05	III	Yes	Yes	FOM	Placebo	Yes	Yes	No	24 weeks	39	No
Sebelipase alfa	LAL-CL02	III	Yes	NR	Biomarker	Placebo	Yes	Yes	Yes	20 weeks	66	Yes
	LAL-CL03	NR	Yes	Yes	Clinical	No	No	n/a	Yes	Up to 208 weeks	9	No
Sapropterin	PKU-003	III	Yes	Yes	Surrogate <sup>a</sup>	Placebo	Yes	Yes	Yes	6 weeks	88	No
	PKU-006	III	Yes	Yes	Surrogate <sup>a</sup>	Placebo	Yes	Yes	Yes	10 weeks	45	No
Betaine	Case reports	n/a	n/a	n/a	Biomarker	n/a	n/a	n/a	Yes	n/a	±140	No
Mercaptopamine	RP103-03	III	Yes	Yes	Surrogate	Active comp.	Yes	No, but justified	Yes	8 weeks	43	Yes
Nitisinone	NTBC study	NR	Yes	Yes	Clinical	No	No	n/a	Yes	NR	207	No
Alipogene	AMT-011-01	NR	No	No	Biomarker	No	No	n/a	Yes	12 weeks	14	No
Tiparovec	AMT-011-02	NR	Yes	No	Biomarker	No	No	n/a	Yes	18 weeks	5	Yes
	AMT-010-01	NR	Yes	Yes	Biomarker	No	No	n/a	Yes	3 months	8	No

Table 3 continued

Drug	Study	Study phase	Multicentre	Multinational	Endpoint	Control	Randomized	Blinding	Study population representative	Duration	No. of patients	QoL-end point
Carglumic acid (organic acidurias)	OE-CCA001-OA2009	IIIb	Yes	Yes	Biomarker	No	No	n/a	Yes	15 days	77	No
Carglumic acid (NAGS deficiency)	Case reports	n/a	n/a	n/a	Biomarker	n/a	n/a	n/a	Yes	n/a	20	No
Zinc	Existing literature	NR	No	No	Biomarker	No	No	n/a	Yes	±3 years	148	No
Cholic acid (Kolbam)	CAC-91-10-10	III	No	No	Biomarker	No	No	n/a	Yes	Up to 17 years	52	No
Cholic acid (Orphacol)	Case reports	n/a	n/a	n/a	Biomarker	n/a	n/a	n/a	Yes	n/a	49	No
Hydrocortisone	DC 06/02	II/III	Yes	No	Biomarker	Active comp.	Yes	No, but justified	Yes	24 weeks	64	Yes
Pasireotide (Cushing)	B2305	III	Yes	NR	Biomarker	Diff. dose	Yes	Yes	Yes	12 months	162	Yes
Pasireotide	C2305	III	Yes	Yes	Biomarker	Active comp.	Yes	Yes	Yes	12 months	358	Yes
(Acromegaly)	C2402	III	Yes	Yes	Biomarker	Active comp.	Yes	Yes	Yes	24 weeks	198	Yes
Ketoconazole	Literature data	n/a	n/a	n/a	Biomarker	n/a	n/a	n/a	Yes	n/a	>800	n/a
Pegvisomant	SEN-3614	III	Yes	Yes	Biomarker	Placebo	Yes	Yes	Yes	12 weeks	112	Yes
Glycerol phenylbutyrate	HPN-100-006	III	NR	NR	Biomarker	Active comp.	Yes	Yes	Yes	4 weeks	44	No
Asfotase alfa	ENB-006-09	II	Yes	Yes	Clinical	Hist. control + diff. dose	Yes	No	Yes	24 weeks	13	Yes
Afamelanotide	ENB-008-10	II	Yes	Yes	Clinical	Hist. control + diff. dose	Yes	No	Yes	Ongoing (extension)	12	Yes
Afamelanotide	CUV039	III	Yes	No	Clinical	Placebo	Yes	Yes	Yes	6 months	93	Yes

FOM functional outcome measure, n/a not applicable, NR not reported, NPC Niemann-Pick C

<sup>a</sup> Endpoint is either (1) validated as a surrogate endpoint in scientific studies, or (2) widely accepted by clinical experts as a valid surrogate for a clinical endpoint



remaining POs either did not want to participate ( $N = 1$ ) or did not reply to e-mails ( $N = 2$ ). On average 19 (range 4–40) patients per participating PO filled in the online survey resulting in a total of 75 patients representing 14 diseases and 13 OMPs. Forty-two patients (56%) reported fatigue and 15 (20%) mentioned pain as the most inconvenient symptom of their disease. When asked for which symptom of their disease patients preferred to see improvement upon treatment with the OMP, 29 (39%) mentioned fatigue and nine (12%) mentioned deterioration of their vital organs or cognition. In contrast, experts seemed to have a preference for endpoints that are convenient to measure, e.g. by blood tests or function tests. A comparison of endpoints used pre- and post-marketing, and preferred endpoints of patients and experts can be found in Supplementary Table 4.

### 3.6 COMPASS Variables

Ten OMPs were excluded from the analysis of factors contributing to the efficacy-effectiveness gap since no post-marketing evidence was available, resulting in a total of 21 OMPs (Table 4). Regarding the type of primary endpoint, 5/7 OMPs (71%) that were authorized based on a study with a clinical or surrogate primary endpoint showed good effectiveness in the real world, versus 3/14 (21%) OMPs that were authorized based on a study with a biomarker or FOM as primary endpoint (Fisher's exact test:  $p = 0.056$ ). Of the OMPs for which the study population was an extreme selection of the patient population, none showed good effectiveness in the real world (category 3), compared to 47% of the OMPs for which the study population was representative for the patient population ( $p = 0.131$ ). Also, 5/16 OMPs (31%) that were authorized for ultra-rare orphan diseases (prevalence of less than 1 in 50,000) had good real-world effectiveness, compared to 3/5 OMPs (60%) that were authorized for rare diseases (prevalence between 1 in 50,000 and 5 in 10,000) ( $p = 0.325$ ). Of the OMPs that were granted a full marketing authorization, 50% had good real-world effectiveness, compared to 22% of the OMPs that were granted authorization under exceptional circumstances ( $p = 0.367$ ). The other COMPASS variables (study phase and randomization) did not show a relevant association with the real-world effectiveness (percentage difference  $\leq 15\%$ , data not shown).

### 3.7 Relationship Between Efficacy and Effectiveness

In total, eight out of 21 authorized OMPs (38%) showed good effectiveness in the real-world situation, and in 16/21 OMPs (76%) the efficacy category corresponded with the real-world effectiveness category. Five out of eight OMPs (63%) for which the efficacy was judged as 'good' showed

good *relative* effectiveness versus 2/12 (17%) of OMPs with no or unclear efficacy (Supplementary Table 5). With respect to the *real-world* effectiveness: five out of seven OMPs (71%) for which the efficacy was judged as 'good' showed good effectiveness versus three out of 14 (21%) with no or unclear efficacy (Supplementary Table 5).

### 3.8 Relationship Between Relative Effectiveness and Real-World Effectiveness

For 13 OMPs no relative effectiveness assessment or post-marketing evidence was available, resulting in a total of 18 OMPs for this analysis. Six out of seven (86%) OMPs for which relative effectiveness was good also showed good real-world effectiveness, versus 1/11 (9%) OMPs with no or unclear relative effectiveness (Supplementary Table 5).

## 4 Discussion

Bridging the gap between efficacy and effectiveness is difficult, particularly in the field of orphan diseases, where generation of evidence is often limited to a few studies with considerable methodological shortcomings. This study reveals that less than half of the approved OMPs for which sufficient post-marketing evidence is available to judge the real-world effectiveness show good effectiveness in the real-world situation. The exploratory analyses of factors contributing to this gap showed that the type of primary endpoint used in the pivotal study seems to be the most important factor. Additional important findings are that none of the OMPs for which the study population was an extreme selection of the patient population showed good effectiveness in the real world, and that a very low disease prevalence and conditional/exceptional authorization also more often led to disappointing real-world effectiveness. Relative effectiveness (at the time of the reimbursement decision) has been shown to be highly correlated with real-world effectiveness. For this study, we chose to compare efficacy assessment from a centralized procedure with relative effectiveness assessment of a national reimbursement authority, which may not represent the situation in all EU countries. However, several general assumptions can be made.

### 4.1 Study Design, Choice of Endpoints and Study Population

Due to the heterogeneity and slowly progressive nature of many rare diseases, there is often a need to use biomarkers or FOMs as the primary study endpoint. Our study, however, suggests that improvements on a biomarker level may not correspond with improvements relevant for the patient

**Table 4** Relationship between COMPASS variables and real-world effectiveness of orphan medicinal products (OMPs)

COMPASS variable	Real-world effectiveness			Fisher's exact
	Good effect (category 3)	Unclear or no clinical effect (categories 0–2)	Total	
<b>Type of endpoint</b>				
Clinical or validated surrogate	5	2	7	
Biomarker or FOM	3	11	14	
Total	8	13	21*	$p = 0.056$
<b>Study population</b>				
Representative for patient population	8	9	17	
Not representative for patient population	0	4	4	
Total	8	13	21*	$p = 0.131$
<b>Disease prevalence</b>				
Rare	3	2	5	
Ultra-rare	5	11	16	
Total	8	13	21*	$p = 0.325$
<b>Type of authorization</b>				
Full	6	6	12	
Exceptional/conditional	2	7	9	
Total	8	13	21*	$p = 0.367$

FOM functional outcome measure

\* OMPs without post-marketing studies were excluded from the analysis: eliglustat, elosulfase alfa, sebelipase alfa, betaine, alipogene tiparovec, cholic acid (Kolbam and Orphacol), glycerol phenylbutyrate, asfotase alfa, afamelanotide

in the real-world setting. According to regulatory guidelines, biomarkers or FOMs may be acceptable but only if validated, and thus labeled as surrogate endpoints. In this respect, it is a common misconception that if an outcome is believed to be correlated to clinical outcomes it can be used as a surrogate endpoint [16]. In-depth studies are needed to confirm that the surrogate endpoint responds to treatment, predicts clinical response and is related to the pathophysiology of the disease. When, at the time of marketing authorization, uncertainty exists about the relationship between the biomarker or FOM and clinical benefit, adequate post-marketing studies and registries may ideally be used to validate the endpoint. Our study shows that if an OMP is authorized based on a relevant effect on a surrogate endpoint, the real-world effectiveness is generally good. This is for example the case in phenylalanine levels in phenylketonuria [17] and platelet counts or spleen volume in Gaucher disease, for which the correlation with clinical outcomes is well established. Interestingly, two of the three OMPs that were approved on the basis of an improvement in biomarker levels but showed good effectiveness in the real-world situation used biomarkers that were—according to the experts—highly correlated with clinical symptoms. Experts' opinions might thus be valuable in judging whether or not a specific endpoint could be used as a surrogate endpoint. Another important note to make is that if a

biomarker or FOM is validated as a surrogate endpoint in a certain disease, this does not mean that it can automatically be used as surrogate endpoint in other diseases. The 6-MWT is for example validated as a surrogate endpoint in pulmonary arterial hypertension, but not for many other diseases. Another important limitation of using the 6-MWT as a primary endpoint is that only patients who are able to walk can be included in the clinical study. Consequently, extrapolating clinical study results to the entire patient population may be problematic [18]. As was shown by our study, none of the OMPs for which the study population was an extreme selection of the patient population showed good effectiveness in the real world. This highlights the need for studies with study populations that are a better reflection of the total patient population, thus having a high external validity.

#### 4.2 Recommendations and Ways Forward

Improvements can be introduced at the pre-marketing and post-marketing stages. First, the quality of pre-marketing studies has to be sufficient to allow an unbiased (independent) judgement about the efficacy of an OMP at the time of marketing authorization. If biomarkers are used as the primary endpoint, caution is required for their predictive value for real-world effectiveness, and studies on long-

term effectiveness are of great importance. Moreover, validation of biomarkers or FOMs as a surrogate endpoint in the context of a specific disease seems appropriate. Since our results suggest that patients' and experts' preferences regarding study endpoints may differ, it is important that all relevant stakeholders (patients, academia, industry and regulatory authorities) agree upon what constitutes a sensitive and validated endpoint early in the development process. Also, selection of the study population and the definition of what constitutes a minimal clinically important difference (MCID) require close attention. Such discussion may take place in the context of a validation procedure at the EMA or a protocol assistance.

Second, since 50% of the OMPs with a full marketing authorization have good real-world effectiveness (in contrast to 22% of OMPs with exceptional circumstances authorization), this underlines the need for generation of coordinated, robust post-marketing evidence. Post-marketing commitments currently often imply the set-up of an industry-sponsored drug registry, which has several shortcomings [19]. Moreover, access to real-world data across the EU is hampered by the lack or inefficiency of cross-border collaborations, fragmentation of resources and the lack of interoperability, which complicates decision making by health technology assessment (HTA) bodies. Hence, there is a need for the implementation of methods to integrate and analyze heterogeneous data. The concept of a self-regulatory market, in which an OMP will automatically no longer be prescribed by doctors when effectiveness is not convincing, might fail in rare diseases with an unmet medical need.

To accelerate access to new drugs, new pathways are increasingly being explored in the EU. The 'adaptive pathway', introduced by the EMA in 2015, enhances timely access of an OMP by approving it in a well-defined patient subgroup with a high (unmet) medical need, followed by widening of the indication to a larger patient population [20]. In addition, HTAs are involved early in the process. Again, the launch of improved disease registries, which can generate robust and independent evidence, is of critical importance [19]. Through this, the importance of timely access is balanced with the need for adequate information on effectiveness and safety, rendering marketing authorization a continuous process [21]. An evaluation on the adaptive pathways pilot showed that this procedure can promote multi-stakeholder dialogues and support development of drugs for which generation of evidence is difficult. Another promising development is the 'real-world evidence' initiative that is currently being developed by the EMA, creating a framework that delivers access to and analysis of multinational real-world data to optimize decision making on medicinal products developed for areas with a high unmet medical need [22].

### 4.3 Limitations

Despite the small sample, we believe that the results of our study may serve as a lead for future studies. Our study is the first to systematically compare pre-marketing data with post-marketing data for OMPs. Studies about the quality of pivotal studies of orphan drugs have been published before [23–25], but they lack the inclusion of post-marketing effectiveness. Since the COMPASS tool is not developed to score or rank the quality of clinical evidence, it was not possible to quantify the level of pre-marketing evidence. Grading of Recommendations Assessment, Development and Evaluation (GRADE) is not appropriate to assess clinical studies of OMPs, as this approach does not take into consideration the methodological drawbacks of these studies [15]. To our knowledge, no other quality assessment tools exist for evaluation of OMPs. Although clinical experts were asked to complete our list of post-marketing studies with studies that they believed were of value for the evaluation of post-marketing effectiveness, we may have missed some studies in the evaluation of post-marketing data. Moreover, since results of negative studies might not be available due to publication bias, the real-world effectiveness as assessed in our study may be an overestimation [26].

## 5 Conclusion

We showed that less than half of the authorized OMPs show good effectiveness in the real-world situation and that the most important contributor to the efficacy-effectiveness gap seemed to be the use of a biomarker in the pivotal study. Whether the results of our research on medicinal products for metabolic orphan diseases can be extrapolated to other orphan disease fields (e.g. oncology, neurology), remains to be elucidated.

**Acknowledgements** We thank all participating clinical experts for sharing their opinions, and Angèl Link for providing the Dutch National Health Care Institute assessments. We also thank the chairpersons of patient organizations for approaching patients, and all patients for filling in the survey.

**Author contributions** CH, MB and YS contributed to the study design. YS was responsible for collecting research data, and interviewing experts and patients. Data interpretation and statistical analysis was performed by YS and supervised by MB. MB, CH, CG and VS critically reviewed the manuscript. All authors read and approved the final manuscript.

### Compliance with Ethical Standards

**Competing interests** YS has no conflicts of interest. Over the past 3 years, MB has received travel support from Sanofi Genzyme once. MB and CH have not performed any paid consultancies nor received

fees for speaking at conferences. The Academic Medical Center participates in clinical trials with Sanofi Genzyme and Protalix and receives support for registries from Sanofi Genzyme and Shire HGT. None of the support or grants are related to the submitted work. This publication is a summary of the European Public Assessment Reports and the summary of product characteristics available on the EMA website, focusing on the main issues discussed during the scientific evaluation. Healthcare professionals and interested readers are referred to the EMA website for up-to-date information on this marketing authorization ([http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/landing/epar\\_search.jsp&mid=WC0b01ac058001d124](http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/landing/epar_search.jsp&mid=WC0b01ac058001d124)). The authors remain solely responsible for the opinions expressed in this publication.

**Funding** This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

**Ethical approval** Not required.

**Data sharing** Data are available from the lead author on request.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Orphanet. The portal for rare diseases and orphan drugs. <http://www.orpha.net/consor/cgi-bin/index.php?lng=EN>. Accessed May 2016.
- Orphan incentives. European Medicines Agency, London. [http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general\\_content\\_000393.jsp](http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000393.jsp). Accessed May 2016.
- EURORDIS Therapeutic Report-Update on therapeutic development and patient involvement in EMA activities. July and August 2016.
- EMA Annual Report 2014. London.
- Pakketbeheer Weesgeneesmiddelen. 26 October 2015.
- Gerss JW, Kopcke W. Clinical trials and rare diseases. *Adv Exp Med Biol.* 2010;686:173–90. doi:10.1007/978-90-481-9485-8\_11.
- Tambuyzer E. Rare diseases, orphan drugs and their regulation: questions and misconceptions. *Nat Rev Drug Discov.* 2010;9(12):921–9. doi:10.1038/nrd3275.
- Wilcken B. Rare diseases and the assessment of intervention: what sorts of clinical trials can we use? *J Inher Metab Dis.* 2001;24(2):291–8.
- Guideline on procedures for the granting of a marketing authorisation under exceptional circumstances, pursuant to article 14 (8) of regulation (EC) No 726/2004 London: European Medicines Agency 2005.
- Vimizim. European Medicines Agency, London. 2014. [http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/002779/smops/Positive/human\\_smop\\_000650.jsp&mid=WC0b01ac058001d127](http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/002779/smops/Positive/human_smop_000650.jsp&mid=WC0b01ac058001d127). Accessed May 2016.
- Eichler HG, Abadie E, Breckenridge A, Flamion B, Gustafsson LL, Leufkens H, et al. Bridging the efficacy-effectiveness gap: a regulator's perspective on addressing variability of drug response. *Nat Rev Drug Discov.* 2011;10(7):495–506. doi:10.1038/nrd3501.
- Group BDW. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther.* 2001;69(3):89–95. doi:10.1067/mcp.2001.113989.
- Straub V, Balabanov P, Bushby K, Ensinì M, Goemans N, De Luca A, et al. Stakeholder cooperation to overcome challenges in orphan medicine development: the example of Duchenne muscular dystrophy. *Lancet Neurol.* 2016;15(8):882–90. doi:10.1016/S1474-4422(16)30035-7.
- Community Register of orphan medicinal products for human use. European Commission. 2016. [http://ec.europa.eu/health/documents/community-register/html/index\\_en.htm](http://ec.europa.eu/health/documents/community-register/html/index_en.htm).
- Picavet E, Cassiman D, Aertgeerts B, Simoens S. Development and validation of COMPASS: clinical evidence of orphan medicinal products—an assessment tool. *Orphanet J Rare Dis.* 2013;8:157. doi:10.1186/1750-1172-8-157.
- Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? *Ann Intern Med.* 1996;125(7):605–13.
- Waisbren SE, Noel K, Fahrback K, Cella C, Frame D, Dorenbaum A, et al. Phenylalanine blood levels and clinical outcomes in phenylketonuria: a systematic literature review and meta-analysis. *Mol Gene Metab.* 2007;92(1–2):63–70. doi:10.1016/j.ymgme.2007.05.006.
- Haas M, Vlcek V, Balabanov P, Salmonson T, Bakchine S, Markey G, et al. European Medicines Agency review of ataluren for the treatment of ambulant patients aged 5 years and older with Duchenne muscular dystrophy resulting from a nonsense mutation in the dystrophin gene. *Neuromuscul Disord NMD.* 2015;25(1):5–13. doi:10.1016/j.nmd.2014.11.011.
- Hollak CE, Aerts JM, Ayme S, Manuel J. Limitations of drug registries to evaluate orphan medicinal products for the treatment of lysosomal storage disorders. *Orphanet J Rare Dis.* 2011;6:16. doi:10.1186/1750-1172-6-16.
- Adaptive pathways. European Medicines Agency. 2016. [http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general\\_content\\_000601.jsp](http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000601.jsp).
- Liu Y, Lacombe D, Stupp R. The changing world of drug development: an academic research organization's perspective on the “Seven Wonders” of the future world of anticancer drug development. *Chin Clin Oncol.* 2014;3(2):19. doi:10.3978/j.issn.2304-3865.2014.05.06.
- Update on Real World Evidence Data Collection. European Medicines Agency 10 March 2016.
- Onakpoya IJ, Spencer EA, Thompson MJ, Heneghan CJ. Effectiveness, safety and costs of orphan drugs: an evidence-based review. *BMJ Open.* 2015;5(6):e007199. doi:10.1136/bmjopen-2014-007199.
- Picavet E, Cassiman D, Hollak CE, Maertens JA, Simoens S. Clinical evidence for orphan medicinal products—a cause for concern? *Orphanet J Rare Dis.* 2013;8:164. doi:10.1186/1750-1172-8-164.
- Putzeist M, Heemstra HE, Garcia JL, Mantel-Teeuwisse AK, Gispén-De Wied CC, Hoes AW, et al. Determinants for successful marketing authorisation of orphan medicinal products in the EU. *Drug Discov Today.* 2012;17(7–8):352–8. doi:10.1016/j.drudis.2011.10.027.
- Mattila T, Stoyanova V, Elferink A, Gispén-de Wied C, de Boer A, Wohlfarth T. Insomnia medication: do published studies reflect the complete picture of efficacy and safety? *Eur Neuropsychopharmacol.* 2011;21(7):500–7. doi:10.1016/j.euroneuro.2010.10.005.