



ScienceDirect

Contents lists available at [sciencedirect.com](http://sciencedirect.com)  
Journal homepage: [www.elsevier.com/locate/jval](http://www.elsevier.com/locate/jval)

## Intraindividual Comparisons to Determine Comparative Effectiveness: Their Relevance for G-BA's Health Technology Assessments

Julia Annabel Wagle, MD, Jan-Paul Flacke, MD, MBA, Dietrich Knoerzer, PhD, Jörg Ruof, MD, MPH, MBA, Sonja Merkesdal, MD, PhD

### ABSTRACT

**Objectives:** Health technology assessments (HTA) rely on head-to-head comparisons. We searched for intraindividual comparisons (IIC) qualifying as head-to-head design to develop comparative evidence.

**Methods:** Gemeinsamer Bundesausschuss (G-BA) appraisals between January 2011 and April 2020 were reviewed for inclusion of IIC. Identified IIC were grouped according to disease characteristics into nonprogressive, progressive, irregular, or symmetrical conditions. Evaluation of IIC by Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) and acceptance of IIC by G-BA were determined, and criteria for the usage and quality of IIC were developed.

**Results:** A total of 483 appraisals finalized between January 2011 and April 2020 were reviewed. Eleven appraisals included IIC: nonacog beta (hemophilia B), turoctocog alpha (hemophilia A), emicizumab (2 appraisals: hemophilia A), pasireotide (unresectable pituitary tumor), lomitapid (homozygous familial hypercholesterolemia), glycerol phenylbutyrate (2 appraisals: urea cycle disorders), asfotase alfa (hypophosphatasia), lumacaftor (cystic fibrosis), and larotrectinib (NTRK<sup>+</sup> solid tumors). All those appraisals related to rare genetic conditions with hemophilia and its bleeding rate are considered mainly a nonprogressive condition. All the other diseases show progressive disease characteristics. None of the identified IIC has been accepted by G-BA. Inconsistencies of before/after study design, lack of clarity on treatments prior to the switch, and different time intervals were among the most commonly cited methodological concerns.

**Conclusions:** IICs provide a rare opportunity to determine comparative effectiveness in distinct clinical settings that are not suitable or difficult to randomize into parallel groups. While manufacturers and researchers should aim for highest methodological standards when running an IIC, HTA bodies should accept IIC in distinct settings when determining relative effectiveness.

**Keywords:** comparative effectiveness, Federal Joint Committee, G-BA, health technology assessment, intraindividual comparison.

VALUE HEALTH. 2021; ■(■):■-■

### Introduction

Innovation in clinical medicine is rapidly changing traditional treatment pathways. Personalized medicines are on the rise, and genetic sequencing and profound molecular insights are allowing ever more targeted therapeutic interventions. Clinical development programs also should explore new and innovative methodologies to determine comparative effectiveness and safety within clinical development programs adapting to these evolving needs.

Randomization is the most established approach in clinical trial design to prevent biases due to structural inhomogeneity of the treatment groups. Its ultimate goal is to enable a comparison of the outcomes of treatments given to groups of patients who do not differ in any systematic way.<sup>1</sup> However, randomized controlled

trials (RCTs) have well-known shortcomings. Most importantly, in rare diseases and targeted therapy regimes, patient samples may be too small for randomization. Thus, the numbers needed to reach structural homogeneity are not achieved. To overcome this limitation, structural homogeneity in clinical conditions that are not suitable for randomization into parallel groups can be achieved by leveraging patients as their own controls. Although those intraindividual comparisons (IIC) are a standard methodology when comparing diagnostic technologies,<sup>2-4</sup> for example, they are also being used in comparative effectiveness research. Intraindividual comparative designs are applied when simultaneously comparing topical treatment outcomes in symmetrical conditions<sup>5-10</sup> or—over time—in before/after clinical study designs.<sup>11,12</sup>

Almost all health technology assessments (HTAs) are based on comparative effectiveness research. While randomized clinical

trials are the base for many of those comparisons, certain conditions qualify for an IIC as state-of-the-art study design to develop the required comparative clinical data. Within the German HTA process Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) conducts the initial assessment of the manufacturers' dossiers, while the Gemeinsamer Bundesausschuss (G-BA) is responsible for the final appraisal of the additional benefit. Thus, we reviewed manufacturers' dossiers, respective IQWiG assessments, and final appraisals by G-BA and determined to what extent IIC were leveraged within G-BA's benefit appraisals.

## Methods

Three steps were conducted to identify the IIC, categorize the underlying disease area, and analyze the IIC, their evaluation by IQWiG, and their acceptance by G-BA (Fig. 1).

### Step 1: Identification of IIC

G-BA's homepage (<https://www.g-ba.de>) was used to identify all benefit appraisals finalized between January 2011 and April 2020. Appraisals were reviewed and screened for inclusion of IIC within either the manufacturer dossier or G-BA's appraisal. Screening was conducted based on Module 1 of the manufacturer dossier and the decision rationale (Tragende Gründe) provided by the G-BA.

We determined the evidence that G-BA's appraisals were based on RCTs or any other, lower-ranked evidence that was provided within the manufacturer's dossiers. In procedures that relied on evidence other than RCTs, we leveraged all documents provided on G-BA's homepage and specifically determined whether IICs (eg, before/after comparisons, within-patient comparisons, intraindividual comparisons) were part of the evidence provided.

### Step 2: Categorization of Disease Conditions

Conditions covered by the IIC were identified and categorized according to the underlying disease characteristics. In particular, 4 disease archetypes were discriminated:

- (1) *Nonprogressive conditions requiring continuous treatment*: a disease characterized by continuous, not progressing signs and symptoms that, however, requires continuous intervention. Those conditions were considered ideal candidates for IIC because they would allow for a head-to-head comparison of structurally identical patient samples (scenario 1, Fig. 2)
- (2) *Progressive conditions*: a disease with a progressive natural course of the disease. The majority of diseases, including all hematology/oncology conditions, diabetes, and cardiovascular disease, show a progressive natural course of the disease. While an IIC seems possible in those conditions, the underlying disease progression would impact comparability of intraindividual time series. The resulting bias might be conservative considering the aggravating character of any disease progression when comparing a later time interval to an earlier timespan intraindividually (ie, comparing B vs A in scenario 2, Fig. 2).
- (3) *Irregular course of disease*: Conditions with a progression that might include temporary remissions as well as repeated relapses (scenario 3, Fig. 2). Multiple sclerosis or various auto-inflammatory conditions might show irregular disease progression characterized by repeated clinical flares and subsequent remissions. Those conditions are considered not suitable for an IIC.

- (4) *Symmetrical conditions*: conditions such as ophthalmologic conditions, for example, that might allow for a simultaneous assessment of investigative and comparative treatment if medicines are applied locally and not systemically.

G-BA appraisals including IIC were scrutinized by 2 authors (SM, JR). Diseases were independently categorized into the 4 scenarios by the 2 authors and finally consolidated among all authors.

### Step 3: Analysis of IIC and Its Acceptance with IQWiG and G-BA

For all identified IIC we analyzed the manufacturer's submission and G-BA's acceptance of the provided evidence for the determination of comparative effectiveness. In particular we reviewed:

- (1) Details of the IIC's study design such as sequential versus simultaneous or solely prospective versus retrospective and prospective intervals
- (2) Details of the provided data including endpoints and outcomes
- (3) Whether the IIC were embedded in a broader conceptualization of comparative effectiveness, that is, whether additional comparative data elements were used to allow for a more comprehensive assessment (eg, overall comparison consisting of different types of studies such as IIC supported or complemented by RCT or indirect comparison)
- (4) Details of IQWiG's scientific assessment of IIC (in cases of orphan designation the scientific assessment is conducted by G-BA)
- (5) Acceptance of IIC within the final appraisal by G-BA.

Finally, based on those findings, criteria were developed to determine suitability and quality of IIC.

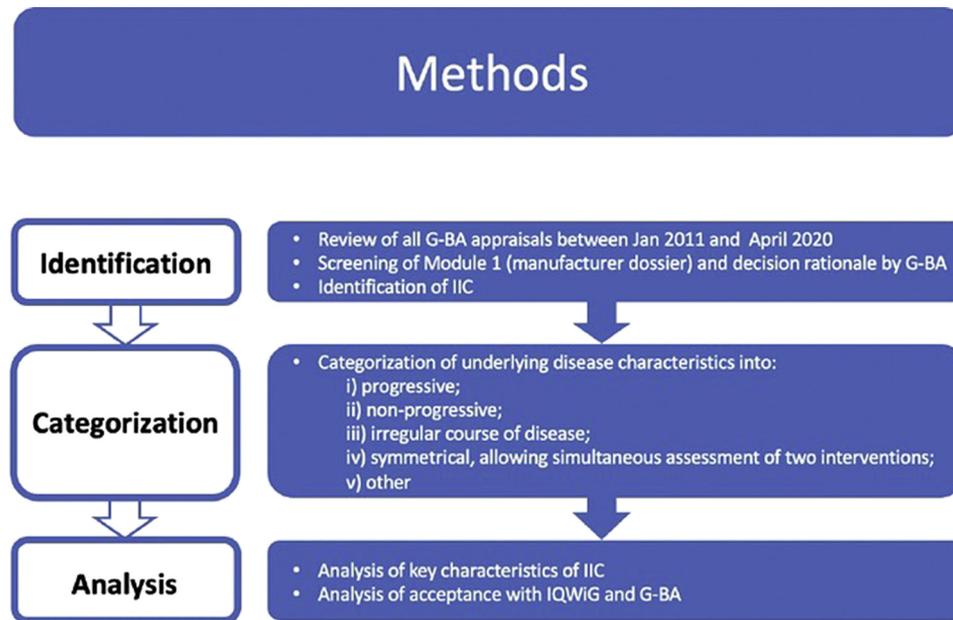
## Results

### Identification of IIC

A total of 483 G-BA appraisals that were finalized between January 2011 and April 2020 were included in our analysis. The most frequent disease areas included oncology/hematology (193 appraisals), metabolic conditions (87 appraisals), and infectious conditions (50 appraisals). Among the 483 appraisals,  $n = 12$  were terminated or put on hold for various reasons;  $n = 20$  appraisals relied on incomplete or missing data because no or only incomplete dossiers had been submitted by the manufacturer; and in  $n = 9$  appraisals, dossiers were submitted but no studies suitable for the assessment of comparative effectiveness were identified and included.

Single arm trials provided the best available evidence in 48 out of the 483 appraisals, with another 15 appraisals relying on uncontrolled (eg, parallel group) data. RCTs provided the main source of clinical data for the remaining procedures.

The screening for IIC revealed 11 appraisals that included intrapatient comparative data: emicizumab and turoctocog alpha in hemophilia A with and without FVIII inhibitors,<sup>13-15</sup> nonacog beta in hemophilia B,<sup>16</sup> lomitapid in patients with homozygous familial hypercholesterolaemia,<sup>17</sup> 2 appraisals of glycerol phenylbutyrate for the treatment of urea cycle disorders,<sup>18,19</sup> asfotase alfa for the treatment of hypophosphatasia,<sup>20</sup> lumacaftor for patients with cystic fibrosis,<sup>21</sup> pasireotide in the treatment of non-resectable

**Figure 1.** Search and analysis strategy.

IIC indicates intraindividual comparison; G-BA, Gemeinsamer Bundesausschuss (Federal Joint Committee); IQWiG, Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (Institute for Quality and Efficiency in Health Care).

pituitary tumors,<sup>22</sup> and larotrectinib for patients with solid tumors<sup>23</sup> (Table 1).

### **Categorization of Disease Characteristics**

Hemophilia A and B are rare genetic conditions that are considered to mainly fall into the archetype of nonprogressive conditions requiring continuous prophylaxis and treatment. While the damage that is caused by, for example, joint or muscle bleedings might be progressive over time, spontaneous bleeding rates are not reported to increase over time,<sup>24</sup> making a comparison of intraindividual bleeding rates following different interventional regimes a suitable target for an IIC.

Homozygous familial hypercholesterolemia (FH) and hypophosphatasia were also categorized as progressive conditions. Both are rare genetic conditions. While FH is a rare and usually life-threatening disease characterized by elevated plasma cholesterol levels, extensive xanthomas, and premature and progressive atherosclerotic cardiovascular disease,<sup>25</sup> hypophosphatasia is characterized by defective bone and teeth mineralization and the related deficiency of serum and bone alkaline phosphatase activity.<sup>26</sup>

Urea cycle disorders result from genetic deficiencies within the urea cycle pathway. Severity of clinical symptoms is influenced by the position of the defective protein in the pathway and the severity of the defect.<sup>27</sup> While there is a wide spectrum of clinical manifestations of urea cycle disorders, the related hyperammonemia and its clinical symptoms usually show a progressive character. Although some disease characteristics could also be considered irregular, we categorized the condition as progressive disease due to its severity and the poor survival rates in untreated newborns.

Cystic fibrosis is a rare genetic disease of the cystic fibrosis transmembrane conductance regulator causing dysfunction of all exocrine glandular cells.<sup>28</sup> Although clinical manifestations vary in

severity, chronic airway infections leading to bronchiectasis and the pancreatic insufficiency show continuously progressive characteristics.

Pituitary tumors inducing hypercortisolism are a heterogeneous group of diseases, and the optimal treatment is surgical resection.<sup>29</sup> Typical features of hypercortisolism are abnormal fat distribution, weight gain, metabolic dysfunction, and hypertension, and the increased mortality is explained by resulting cardiovascular disease, diabetes, and infections. This clinical condition is therefore categorized as progressive.

Solid tumors with a neurotrophic tyrosine receptor kinase gene fusion (NTRK) are ultrarare progressive diseases and manifest in various locations independent of underlying histology.<sup>30</sup>

### **Analysis of IIC and Its Acceptance With IQWiG and G-BA**

An overview of identified IIC is provided in Tables 1 and 2. In all 11 HTA processes, the IIC was a key component of comparative evidence provided within the manufacturers' dossiers. Within the appraisals of nonacog beta pegol, turoctocog alpha pegol, lumacaftor, larotrectinib, and the second appraisals each of glycerol phenylbutyrate and emicizumab, the IIC was the only component of comparative evidence the manufacturer relied on in their benefit claim.

The 2 glycerol phenylbutyrate, the asfotase alfa, and the pasireotide appraisals related to an orphan designation with an additional benefit being granted by law. Furthermore, the first appraisal of emicizumab with FVIII inhibitors achieved a hint (lowest evidence category) for a non-quantifiable additional benefit. This benefit was based mainly on the data from the randomized HAVEN 1 trial. Within the first assessment of lomitapid no dossier was submitted, while the second submission included the IIC data.<sup>17</sup> No additional benefit was achieved, and the manufacturer decided to withdraw the medicine from the German market. Nonacog did not achieve an additional benefit.

None of the identified IIC related to a symmetrical condition that would have allowed for a simultaneous assessment of 2

**Table 1.** Appraisals included in our analysis.

Substance (trade name)	Indication (estimated population size)	Date G-BA appraisal	Additional benefit	Characteristic of condition	Evidence base main evidence (supportive evidence)
Emicizumab (Hemlibra)	Hemophilia A, no FVIII inhibitors (n = ~2000)	Sep 5, 2019	No additional benefit	Nonprogressive disease	Intraindividual comparison Partially randomized trial Noninterventional study (Indirect treatment comparison)
Emicizumab (Hemlibra)	Hemophilia A with FVIII inhibitors (n = ~100)	Sep 20, 2018	Additional benefit not quantifiable <sup>†</sup>	Nonprogressive disease	Partially randomized trial Single arm trial Noninterventional study Intraindividual comparison Indirect treatment comparison
Nonacog beta pegol (Refixia)	Hemophilia B (n = 500–570)	Apr 19, 2018	No additional benefit	Nonprogressive disease	Single arm trial Intraindividual comparison
Turoctocog alpha pegol (Esperoct)	Hemophilia A patients ≥12 years (n = 2840–3190)	Feb 6, 2020	No additional benefit	Nonprogressive disease	Multi-armed nonrandomized trial Intraindividual comparison
Lomitapid (Lojuxta)	Homozygous familial hypercholesterolemia (n = 60–70)	Nov 27, 2015	No additional benefit	Progressive disease	Single arm trial Extension study Intraindividual comparison (Registry)
Glycerol phenylbutyrate (Ravicti)*	Urea cycle disorders in newborns 0-2 months (n = 10–18)	Jul 4, 2019	Additional benefit not quantifiable	Progressive disease	Single arm trial Intraindividual comparison
Glycerol phenylbutyrate (Ravicti)*	Urea cycle disorders (n = 100–250)	Aug 16, 2018	Additional benefit not quantifiable	Progressive disease	Randomized controlled trial cross-over extension Single arm trials Intraindividual comparison
Asfotase alfa (Strensiq)*	Hypophosphatasia (n = ~1000)	Mar 17, 2016	Additional benefit not quantifiable	Progressive disease	Randomized controlled trial Single arm trial Historic controls (retrospective observational study) Intraindividual comparison
Lumacaftor (Orkambi)	Cystic fibrosis Children 2-5 years (n = ~280)	Aug 15, 2019	Additional benefit not quantifiable	Progressive disease	Single arm trial Intraindividual comparison
Pasireotide (Signifor)*	M. Cushing in non-resectable pituitary tumor (n = 160–360)	Dec 6, 2012	Additional benefit low	Progressive disease	Randomized two-armed trial Intraindividual comparison
Larotrectinib (Vitrakvi)	NTRK + Solid tumors (n = 390–770)	Apr 2, 2020	No additional benefit	Progressive disease	Single arm trials Intraindividual comparison

\*Orphan designation.

<sup>†</sup>One subgroup only.

**Table 2.** IIC included in our analysis.

Substance (trade name)	IIC Key characteristic	IIC Key efficacy endpoint	IQWiG assessment (in Orphan: G-BA Assessment)	G-BA appraisal comments
Emicizumab (Hemlibra)	Time period before and after emicizumab treatment	Annual bleeding rate	Prior therapy without appropriate prophylaxis; size of effect not sufficient; different settings of data collection prior versus after switch	IIC not accepted because of methodological limitations
Emicizumab (Hemlibra)	Time period before and after emicizumab	Annual bleeding rate	Prior therapy without appropriate prophylaxis	IIC not accepted because of methodological limitations: different settings of data collection, ie, uncontrolled (prior to switch) versus controlled setting (after switch); unexplained dropouts; lack of information regarding duration of observation period
Nonacog beta pegol (Refixia)	Time period before and after nonacog	Annual bleeding rate	Data not sufficient to determine additional benefit	IIC not accepted because of lack of clarity regarding prior therapy; different settings of data collection, ie, non-trial (prior to switch) vs trial setting (after switch)
Turoctocog alpha pegol (Esperoct)	Time period before and after turoctocog treatment	Annual bleeding rate	No comparability of settings prior and after therapy, only selected outcomes, risk of bias in IIC too high	IIC not accepted because of methodological limitations
Lomitapid (Lojuxta)	LDL-C level at baseline vs after switch	LDL-C levels	IIC considered weak evidence; Provided evidence incomplete; Endpoint not validated; Questions regarding prior treatment; Lack of appropriate adverse event analysis	IIC not accepted because of lack of information regarding LDL lowering therapy prior to switch; suboptimal LDL-apheresis frequency; inconsistent measurement of LDL-C; incomplete adverse event analysis
Glycerol phenylbutyrate (Ravicti)	Baseline vs follow-up	Successful switch to glycerol phenylbutyrate	Data do not allow to draw conclusions; Number of hyper ammonia crisis unclear	Study and study endpoint not accepted as valid comparative design. 6 out of 16 children dropped out of the study
Glycerol phenylbutyrate (Ravicti)	Retrospective vs prospective and baseline vs follow-up	Plasma ammonia levels; hyper ammoniaic crisis	Plasma ammonia levels not considered relevant for patients; number of hyper ammoniaic crisis too low to determine relative risk; no conclusions can be drawn	IIC not accepted because of recall bias; different time intervals; different assessment of endpoints; lack of valid control

*continued on next page*

**Table 2.** Continued

Substance (trade name)	IIC Key characteristic	IIC Key efficacy endpoint	IQWIG assessment (in Orphan: G-BA Assessment)	G-BA appraisal comments
Asfotase alfa (Strensiq)	Outcomes at baseline vs after switch	Comprehensive test battery incl. size; weight (Z-score); motor function; mobility	IIC have a high risk of bias; no conclusions can be drawn from IIC	All assessments indicate an improvement in motor function; lack of information on minimal clinically relevant difference. General concerns regarding validity of before/after comparisons
Lumacaftor/ivacaftor (Orkambi)	Baseline vs after 24 weeks	Lung Clearance Index	Evidence not sufficient since no direct comparison possible, effects not statistically significant	IIC renders insufficient data, additional benefit based on extrapolation to similar analyses in age group 6-11 and $\geq 12$ years
Pasireotide (Signifor)	Morbidity and SMR at baseline vs after 6 mo	Urine cortisol level	IIC evidence limited, urine cortisol level validated for morbidity assessment, low additional benefit can be stated	Low additional benefit regarding urine cortisol level No additional benefit regarding clinical outcomes and SMR as surrogate for mortality since validity not proven
Larotrectinib (Vitrakvi)	Time period before and after larotrectinib treatment	Time until progression	No evidence on distinct tumor entities given Overall comparison of effectiveness summarizing disease entities render no conclusion on benefit	Comparative design not accepted as valid

IIC indicates intraindividual comparison; G-BA, Gemeinsamer Bundesausschuss; IQWIG, Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen; SMR, standardized mortality ratio.

different topical treatments. Instead, all 11 IIC leveraged sequential assessments over time (Fig. 3). While the 4 hemophilia IIC compared the bleeding rate of 2 different time periods, larotrectinib the time to progression of the time period before and after treatment, the remaining appraisals compared specific measures at baseline and follow-up (ie, assessments were related to time points rather than time intervals).

All appraisals provided various components of evidence with 4 dossiers including also data from randomized or partially randomized parallel group trials. In asfotase alfa the trial ENB-006-09 compared 2 different dosing regimens of asfotase alfa in 6 versus 7 patients and the trial ENB-009 compared no treatment (6 patients) to 2 dosing regimens of asfotase alfa (6 and 7 patients). The pivotal trial of the first appraisal of glycerol phenylbutyrate HPN-100-006 leveraged a 2-week cross-over design to compare glycerol phenylbutyrate to NaPBA (sodium phenylbutyrate) in 22 and 24 patients per group. Finally, the HAVEN 1 trial included 109 hemophilia A patients with FVIII inhibitors and compared bleeding rates of emicizumab versus placebo across 4 different treatment arms,<sup>31</sup> and the HAVEN 3 trial included a total of 152 patients with hemophilia A without FVIII inhibitors across 3 treatment arms.<sup>32</sup>

Key endpoints within the IIC included surrogate parameters such as LDL-C (lomitapid), plasma ammonia levels (glycerol phenylbutyrate), urine cortisol level (pasireotide), functional tests (asfotase alfa), time until disease progression (larotrectinib), lung clearing index (lumacaftor), and annual bleeding rate (hemophilia products). The second appraisal of glycerol phenylbutyrate was

related to newborns 0–2 years of age with urea cycle disorders. Here, the primary study endpoint was the successful switch of treatment. Within the first appraisal of glycerol phenylbutyrate the endpoint reduction of plasma ammonia level was accepted by G-BA. Within G-BA's appraisal of lomitapid the specific measurement of LDL-C (prior to LDL apheresis) was considered biased without commenting on the acceptance of this endpoint more generally.

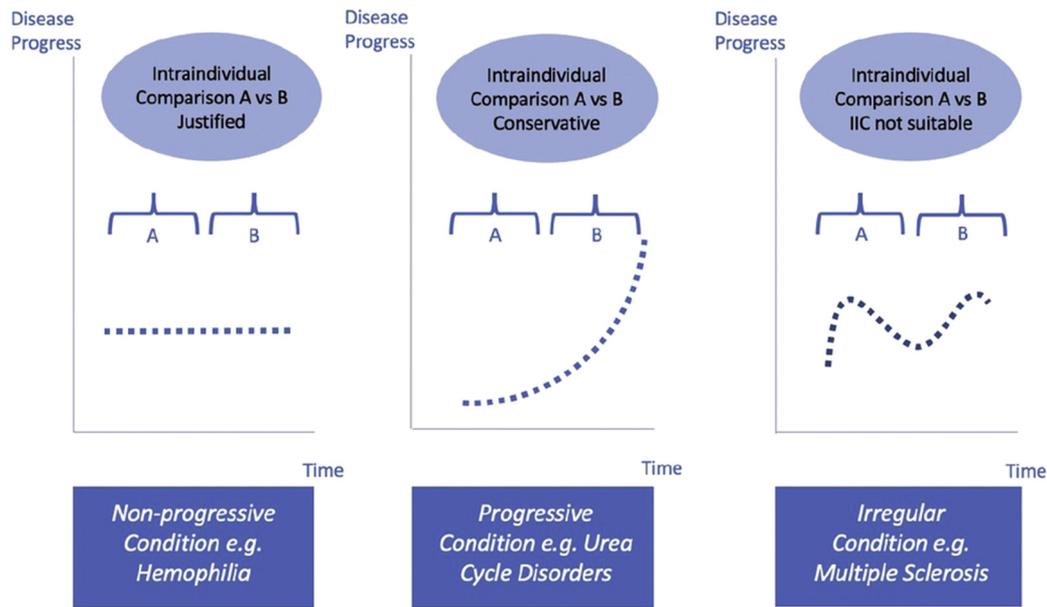
All IIC data were neither evaluated positively by IQWIG nor accepted by G-BA in any of the identified appraisals. Inconsistencies of retrospective/prospective study design, lack of clarity on treatments prior to the switch, and different time intervals were among the most commonly cited methodological concerns (see Table 2). Both IQWIG and G-BA have a clear preference for RCT data and consider any alternative source of evidence to be inferior.

### Development of Checklist on Criteria for IIC Employment

The suggested checklist covering criteria for an IIC employment is included in Table 3. When summarizing the evidence derived from the IIC review, it becomes obvious that the need for an IIC arises when disease characteristics do not allow for a randomization. This typically does occur in (ultra-) rare conditions, or in diseases that display heterogeneous disease characteristics across the individual patients.

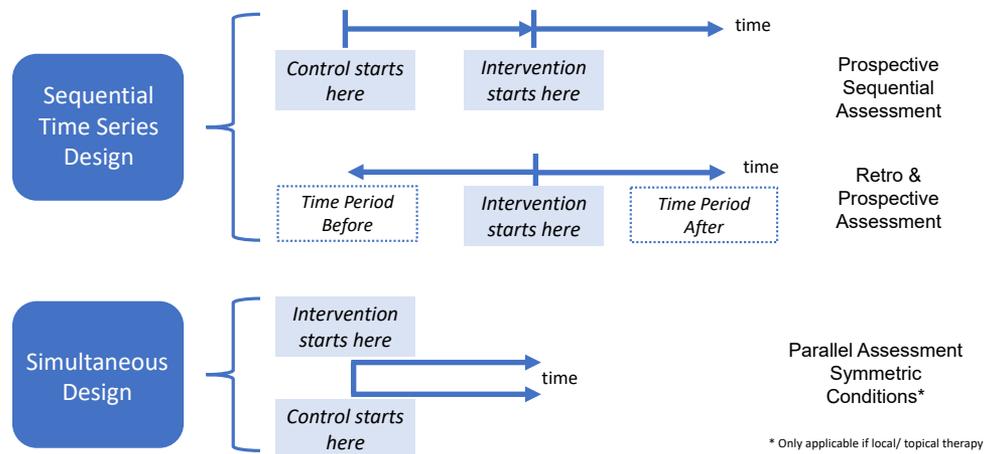
If the overall course of the disease is progressive, the conduct of an IIC with the standard of care treatment being tested first and

**Figure 2.** Disease archetypes and applicability of longitudinal intraindividual study designs (comparing A versus B). Symmetric conditions that might also allow for an intraindividual comparison of locally applied treatments are not covered within this figure.



\* Symmetric conditions that might also allow for an intraindividual comparison of locally applied treatments are not covered within this figure

**Figure 3.** Different types of IIC.



the innovative medicine subsequently does include a systematic bias. As this bias is conservative (ie, against the innovative medicine), the conduct of an IIC should be accepted in that setting.

Finally, any IIC should adhere to best possible methodological standards, that is, control and interventional treatment should be included in one study protocol and all assessments (time periods/endpoints, etc) should be identical.

## Discussion

Due to ongoing progress in biochemical and molecular genetical insights, preclinical research is subject to rapid change and the development of innovative treatments.<sup>33</sup> The integration of targeted therapies in rare conditions in clinical practice is

playing an ever more important role. While parallel group RCTs remain the gold standard to determine comparative effectiveness,<sup>34</sup> alternative head-to-head study designs should also be taken into consideration, in particular in settings where randomization in parallel groups is not feasible or possible. In those situations, IIC may offer certain advantages such as structural identity of control and intervention group, facilitation of patient recruitment in ultrarare conditions or ethical compatibility in severe or life-threatening settings (eg, late/last line oncology treatments) where no effective control therapies are available. Inclusions of such innovative study designs into comparative HTA considerations by IQWiG and G-BA therefore seem timely and reasonable to match preclinical innovation with innovation in clinical research designs. However, our review revealed that:

**Table 3.** Checklist for the suitability and quality of IIC.

Is the condition and intervention difficult/not applicable to randomization, eg, <ul style="list-style-type: none"> <li>• Very rare (eg, genetic or personalized) conditions</li> <li>• Heterogeneous disease characteristics across individual patients</li> <li>• Type of intervention (eg, surgical) difficult to randomize</li> </ul>
Do condition and intervention allow for an IIC <ul style="list-style-type: none"> <li>• Nonprogressive disease</li> <li>• Symmetrical condition with local/topical intervention</li> </ul>
Can intervention and control be covered within one prospective study protocol
Are all assessments (time period/endpoints, etc) identical across control and intervention

- IIC are only applied in 11 out of 483 G-BA appraisals
- G-BA and IQWiG did not accept any of the provided IIC

The G-BA's and IQWiG's view on the use of non-RCT evidence is known to be reluctant. The IQWiG Method paper version 6 (34, chapter 9.1.2) states that "In the first place, a control group is required for the benefit assessment of interventions. From a pure pre-post comparison in a design with dependent samples without a control group, usually no evidence for an effect of an intervention can be derived. Exceptions are clinical conditions with a deterministic (or almost deterministic) course (e.g. diabetic ketoacidotic coma)." But since this group of exceptions might be more relevant than recognized up to now, the present article aims at evaluating this focus of evidence generation more closely.

### Limited Usage of IIC

Because the main strength of IIC is structural identity of the control and intervention group, the limited usage of IIC in clinical development settings is striking. Conceptually, we considered (1) rare nonprogressive conditions with an ongoing need for intervention to control the disease burden to the patients and (2) symmetric conditions in case the innovative and control treatments are applied locally as best candidates for an IIC. Hemophilia was a condition considered to best represent the first archetype. Randomization is not a standard technology in hemophilia clinical research. Baseline characteristics for the various treatment arms in the 2 partially randomized emicizumab trials HAVEN 1 and HAVEN 3 differed considerably, for example, with regard to age of the patients and overall patient numbers included in the trials were very limited, making this condition an ideal candidate for an IIC to develop comparative efficacy data. Nevertheless, only 3 out of the total of 12 G-BA appraisals in hemophilia products between 2011 and April 2020 included an IIC, clearly indicating that despite the advantages of this study design, the usage within suitable clinical development programs is still limited. Also, none of the assessments of aflibercept in symmetrical ophthalmological conditions such as diabetic retinopathy or neovascular age-related macular degeneration leveraged IIC. Sham injections in those conditions were applied to different patients rather than to the second eye of the same patient within the respective clinical development programs of aflibercept.<sup>35,36</sup>

Five IIC were identified covering genetic metabolic conditions (urea cycle disorders; 2 appraisals) and 1 each in homozygous familial hypercholesterolemia, cystic fibrosis, and hypophosphatasia. While clinical features in those rare conditions might differ considerably, we considered all of them to show

rather progressive disease characteristics. Progressive conditions cause a potential source of bias when conducting an IIC because patients with later line treatments tend to show more disease symptoms than treatment naïve or early line patients. Applying the comparative treatment first (treatment A; scenario 2 in Fig. 2) followed by the innovative treatment after failure of the comparative regimen (treatment B; scenario 2 in Fig. 2) might therefore create a systematic bias in favor of the comparative treatment. Conducting an IIC in progressive conditions therefore seems biased although justified, and more frequent usage of this clinical trial design might be feasible.

### Limited Acceptance of IIC

G-BA's and IQWiG's rationale to not accept any of the provided IIC may indicate unrealistically high hurdles regarding comparative effectiveness in certain conditions. As previously demonstrated in the analysis of indirect treatment comparisons, HTA bodies could benefit by taking into consideration comparative clinical evidence beyond RCTs alone.<sup>37</sup> Questions regarding the prior therapy, different settings of data collection prior (eg, within a registry) and after switch (within a single arm trial protocol), different observation periods, and the relevance of surrogate endpoints to patients all have to be addressed but should not lead to delayed access of patients to helpful innovative treatments. Methodological limitations should always be weighted versus the need to consider best available evidence for each research question to be solved.

- Regulatory bodies often refer to the "totality of evidence"<sup>38</sup> and include any available source of preclinical and clinical evidence in their decision making. IIC are an accepted source of evidence within the regulatory environment. In the assessment of emicizumab EMA suggested that the "intra-patient comparison ... demonstrated a clinically meaningful reduction in rates of treated bleeds and all bleeds for emicizumab prophylaxis compared with before episodic FVIII treatment."<sup>39</sup> Similarly, HTA bodies should take an integrative approach to the various components of comparative effectiveness (ie, accept the totality of comparative evidence rather than almost categorically rejecting key components such as all analyzed IIC).
- While IIC in hemophilia compared bleeding rates of the time intervals before and after switch, the 5 metabolic IIC primarily relied on a comparison of baseline and follow-up scores before and after switch (ie, a comparison of time points rather than time intervals). Comparison of sequential time intervals should be acknowledged as the superior methodology and standards for the quality of IIC as a reliable tool should be defined. IIC are to be planned prospectively according to high methodological requirements.

In conclusion, IICs provide an opportunity to determine comparative effectiveness in distinct clinical settings that are not suitable to randomization into parallel patient groups.<sup>40</sup> While manufacturers and researchers should aim for the highest methodological standards when running an IIC, HTA bodies should integrate IIC into their methodological armamentarium when determining relative effectiveness (Table 3).

### Article and Author Information

**Accepted for Publication:** November 28, 2020

doi: <https://doi.org/10.1016/j.jval.2020.11.016>

**Author Affiliations:** Roche Pharma AG, Grenzach-Wyhlen, Germany (Wagle, Flacke, Knoerzer); Medical School of Hanover, Hanover, Germany and r-connect ltd (Ruof, Merkesdal); r-connect ltd, Basel, Switzerland (Ruof, Merkesdal).

Address correspondence to: Joerg Ruof, MD, MPH, MBA, r-connect ltd, Hauensteinstr. 132, 4059 Basel, Switzerland. Email: [joerg.ruof@r-connect.org](mailto:joerg.ruof@r-connect.org)

**Author Contributions:** *Concept and design:* Wagle, Flacke, Knoerzer, Ruof, Merkesdal

*Acquisition of data:* Flacke, Ruof, Merkesdal

*Analysis and interpretation of data:* Wagle, Flacke, Knoerzer, Ruof, Merkesdal

*Drafting of the manuscript:* Wagle, Flacke, Ruof, Merkesdal

*Critical revision of the paper for important intellectual content:* Wagle, Flacke, Knoerzer, Ruof, Merkesdal

*Statistical analysis:* Knoerzer

*Obtaining funding:* Flacke, Knoerzer, Ruof

*Administrative, technical, or logistic support:* Flacke, Knoerzer

*Supervision:* Knoerzer

**Conflict of Interest Disclosures:** Drs Wagle, Flacke, and Knoerzer are employed by Roche Pharma AG. Drs Ruof and Merkesdal reported receiving grants from Roche Pharma AG during the conduct of the study.

**Funding/Support:** This work was funded by Roche Pharma AG.

**Role of the Funder/Sponsor:** The funder was involved in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

## REFERENCES

- Altman DG, Bland JM. Treatment allocation in controlled trials: why randomise? *BMJ*. 1999;318:1209.
- Park M, Chung YE, Lee HS, et al. Intraindividual comparison of diagnostic performance in patients with hepatic metastasis of full-dose standard and half-dose iterative reconstructions with dual-source abdominal computed tomography. *Invest Radiol*. 2014;49:195–200.
- Yim JH, Kim YK, Min JH, Lee J, Kang TW, Lee SJ. Diagnosis of recurrent HCC: intraindividual comparison of gadolinic acid MRI and extracellular contrast-enhanced MRI.
- Schley G, Köberle C, Manuilova E, et al. Comparative analysis of diagnostic and predictive performance of novel renal biomarkers in plasma and urine of acute kidney injury patients. *Intensive Care Med Experiment*. 2015; 3(Suppl 1):A258.
- Tzelikis PF, Morato CS, Neves NT, Hida WT, Alves MR. Intraindividual comparison of nepafenac 0.3% for the prevention of macular edema after phacoemulsification. *J Cataract Refract Surg*. 2018;44:440–446.
- Gonnermann J, Bertelmann E, Pahlitzsch M, Maier-Wenzel AB, Torun N, Klarman MK. Contralateral eye comparison study in MICS & MIGD: Trabectome vs iStent inject. *Graefes Arch Clin Exp Ophthalmol*. 2017;255:359–365.
- Augustin M, Maier K, Sommer B, Sattler G, Herberger K. Double-blind, randomized, intraindividual comparison study of the efficacy of prilocaine and lidocaine in tumescent local anesthesia. *Dermatology*. 2010;221: 248–252.
- Lecompte F, Vignion-Dewalle AS, Vicentini C, et al. Evaluating the non-inferiority of a new photodynamic therapy (Flexitheralight) compared with conventional treatment for actinic keratosis: protocol for a phase 2 study. *JMIR Res Protoc*. 2019;26.
- Vrani F, Sotiriou E, Lazaridou E, et al. Short incubation fractional CO<sub>2</sub> laser-assisted photodynamic therapy vs. conventional photodynamic therapy in field-cancerized skin: 12-months follow-up results of a randomized intraindividual comparison study. *J Eur Acad Dermatol Venereol*. 2019;33:79–83.
- Dirschka T, Ekanayake-Bohlig S, Dominicus R, et al. A randomized, intraindividual, non-inferiority, Phase III study comparing daylight photodynamic therapy with BF-200 ALA gel and MAL cream for the treatment of actinic keratosis. *J Eur Acad Dermatol Venereol*. 2019;33:288–297.
- Matowe LK, Leister CA, Crivera C, Korth-Bradley JM. Interrupted time series analysis in clinical research. *Ann Pharmacother*. 2003;37:1110–1116.
- Young G, Liesner R, Chang T, et al. A multicenter, open-label phase 3 study of emicizumab prophylaxis in children with hemophilia A with inhibitors. *Blood*. 2019;12:2127–2138.
- Gemeinsamer Bundesausschuss. Nutzenbewertungsverfahren zum Wirkstoff Emicizumab. 2018. <https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/355/>. Accessed February 27, 2020.
- Gemeinsamer Bundesausschuss. Nutzenbewertungsverfahren zum Wirkstoff Emicizumab (neues Anwendungsgebiet). 2019. <https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/449/>. Accessed February 27, 2020.
- Gemeinsamer Bundesausschuss. Nutzenbewertungsverfahren zum Wirkstoff Turoctocog alfa pegol. 2020. <https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/478/>. Accessed August 25, 2020.
- Gemeinsamer Bundesausschuss. Nutzenbewertungsverfahren zum Wirkstoff Nonacog beta pegol. 2018. <https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/325/>. Accessed February 27, 2020.
- Gemeinsamer Bundesausschuss. Nutzenbewertungsverfahren zum Wirkstoff Lomitapid. 2015. <https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/175/>. Accessed February 27, 2020.
- Gemeinsamer Bundesausschuss. Nutzenbewertungsverfahren zum Wirkstoff Glycerol phenylbutyrate. 2018. <https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/351/>. Accessed February 27, 2020.
- Gemeinsamer Bundesausschuss. Nutzenbewertungsverfahren zum Wirkstoff Glycerol phenylbutyrate (neues Anwendungsgebiet). <https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/435/>. Accessed February 27, 2020.
- Gemeinsamer Bundesausschuss. Nutzenbewertungsverfahren zum Wirkstoff Asfotase alfa. 2016. <https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/190/>. Accessed February 27, 2020.
- Gemeinsamer Bundesausschuss. Nutzenbewertungsverfahren zum Wirkstoff Lumacaftor/ Ivacaftor. 2021. <https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/444/>. Accessed August 25, 2020.
- Gemeinsamer Bundesausschuss. Nutzenbewertungsverfahren zum Wirkstoff Pasireotid. 2012. <https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/34/>. Accessed August 25, 2020.
- Gemeinsamer Bundesausschuss. Nutzenbewertungsverfahren zum Wirkstoff Larotrectinib" before the the year 2020. <https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/502/>. Accessed August 25, 2020.
- Risal A, Carpenter SL. Hemophilia A and B: an overview: hospital physician. *Hematology/Oncology*. 2017;12:15–31.
- Cuchel M, Bruckert E, Ginsberg, et al. Homozygous familial hypercholesterolaemia: new insights and guidance for clinicians to improve detection and clinical management. A position paper from the Consensus Panel on Familial Hypercholesterolaemia of the European Atherosclerosis Society. *Eur Heart J*. 2014;35:2146–2157.
- Mornet E. Hypophosphatasia. *Orphanet J Rare Dis*. 2007;2:40.
- Mew NA, Simpson KL, Gropman AL, Lanpher BC, Chapman KA, Summar MK. *Urea Cycle Disorders Overview*. GeneReviews. Seattle (WA): University of Washington, Seattle; 2003.
- Goetz D, Ren CL. Review of cystic fibrosis. *Pediatr Ann*. 2019;48:e154–e161.
- Nieman LK, Ilias I. Evaluation and treatment of Cushing's syndrome. *Am J Med*. 2005;118:1340–1346.
- Vaishnavi A, Le AT, Doeble RC. TRKING down an old oncogene in a new era of targeted therapy. *Cancer Discov*. 2015;5:25–34.
- Oldenburg J, Mahlangu JN, Kim B, et al. Emicizumab prophylaxis in hemophilia A with inhibitors. *NEJM*. 2017;377:809–818.
- Mahlangu J, Oldenburg J, Paz-Priel I, et al. Emicizumab prophylaxis in patients who have hemophilia A without inhibitors. *NEJM*. 2018;379:811–822.
- Haendel MA, Chute CG, Robinson PN. Classification, ontology, and precision medicine. *NEJM*. 2018;379:1452–1462.
- IQWiG. General methods; draft version 6.0. <https://www.iqwig.de/en/methods/methods-paper.3020.html>. Accessed September 10, 2020.
- Gemeinsamer Bundesausschuss. Nutzenbewertungsverfahren zum Wirkstoff Aflibercept. 2015. <https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/50/>. Accessed February 12, 2021.
- Gemeinsamer Bundesausschuss. Nutzenbewertungsverfahren zum Wirkstoff Aflibercept (neues Anwendungsgebiet). 2015. <https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/139/>. Accessed February 12, 2021.
- Werner S, Lechterbeck L, Rasch A, Merkesdal S, Ruof J. Analysis of acceptance rate and rationales for rejection of indirect comparisons in IQWiG's benefit assessments. *Gesundh ökon Qual Manag*. 2020;25:24–36.
- Wolff-Holz E, Tiitso K, Vleminchx C, Weise M. Evolution of the EU biosimilar framework: past and future. *BioDrugs*. 2019;33:621–634.
- European Medicines Agency. Assessment Report Hemlibra. [https://www.ema.europa.eu/en/documents/variation-report/hemlibra-h-c-4406-ii-0002-epar-assessment-report-variation\\_en.pdf](https://www.ema.europa.eu/en/documents/variation-report/hemlibra-h-c-4406-ii-0002-epar-assessment-report-variation_en.pdf). Accessed September 10, 2020.
- Yeh RW, Valsdottir LR, Yeh MW, et al. Parachute use to prevent death and major trauma when jumping from aircraft: randomized controlled trial. *BMJ*. 2018;363:k5094.