Journal of Pharmacokinetics and Pharmacodynamics (2019) 46:499–509
https://doi.org/10.1007/s10928-019-09654-3

ORIGINAL PAPER

A modeling and simulation-based assessment of the impact of confounding factors on the readout of a sildenafil survival trial in pulmonary arterial hypertension

Pascal Chanu[1] · Xiang Gao[2] · Rene Bruno[3] · Laurent Claret[3] · Lutz Harnisch[4]

Received: 7 December 2018 / Accepted: 11 September 2019 / Published online: 20 September 2019

Abstract

Sildenafil (REVATIO®) was approved for the treatment of adult Pulmonary Arterial Hypertension (PAH) in the US and the EU. A pediatric study has been performed and sildenafil was approved in the EU for pediatric PAH. The long-term extension of this study revealed good survival but also an increased mortality with the high dose of sildenafil compared to lower doses. As a consequence, FDA required Pfizer to evaluate REVATIO®'s effect on the risk of death in adults with PAH. Following FDA's rationale a survival model was developed to characterize the exposure–mortality relationship and assess its potential impact on an ongoing survival trial in adults in the context of confounding factors. Clinical trial simulations were performed to assess the design of the survival trial in adults (AFFILIATE, NCT02060487), expected to last approximately 8 years according to both assumptions: absence or presence of an exposure–mortality relationship and to quantify the impact of confounding factors on its readout. Simulations showed that the trial would be robust in most conditions. But its interpretation will depend on the number of confounding factors such as additional treatments attempting to control disease progression.

Clinical trial identifier NCT00159913 for STARTS-1, NCT00159874 for STARTS-2

Keywords Survival · Pulmonary arterial hypertension · Clinical trial simulations · Confounding

Introduction

Sildenafil (REVATIO®), 20 mg given three times a day (tid), received approval for the treatment of adult pulmonary arterial hypertension (PAH) in the US (current FDA drug label January 17, 2019) and the EU (current EMEA product information July 1, 2019). Study STARTS-1 (n = 234, ages 1–17 years, NCT00159913) had been performed in pediatric patients [1] and sildenafil was approved in the EU for the treatment of pediatric PAH, at 10 mg tid for children ≤ 20 kg body weight, and 20 mg tid for children above 20 kg. Approvals were based on exercise capacity outcomes: 6 min walking distance and oxygen consumption (VO2 peak), for adult and pediatric populations, respectively. The development of drugs in PAH is evolving with an increased interest for clinical worsening events, mortality being one of them [2, 3]. The long-term extension of the pediatric study, STARTS-2 (NCT00159874) revealed 3 year survival of 84% and 94% in patients > 20 kg and ≤ 20 kg, respectively [1]. Those results are consistent with previously reported results [4]. Despite good overall survival in STARTS-2, a higher risk of mortality occurred among patients randomized to the high-dose compared with lower doses of sildenafil. [5] Causes of death were typical of patients with PAH. During the study, there were 42 reported deaths, with 37 of these deaths reported prior to a decision to titrate subjects to a

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10928-019-09654-3) contains supplementary material, which is available to authorized users.

✉ Lutz Harnisch
Lutz.O.Harnisch@pfizer.com

[1] Clinical Pharmacology, Genentech/Roche, Lyon, France

[2] Alexion Pharmaceuticals, Pharmacometrics and Physiological PK/PD Modeling and Simulation, Cheshire, CT, USA

[3] Clinical Pharmacology, Genentech/Roche, Marseille, France

[4] Pfizer, Pharmacometrics/Global Clinical Pharmacology, Sandwich, UK

lower dosage because of a finding of increased mortality with increasing REVATIO® doses. A survival analysis was conducted on 37 deaths data cutoff and the hazard ratio (HR) for high dose compared to low dose was 3.9, p = 0.007. Because this long-term extension did not include a placebo arm, the impact of sildenafil monotherapy on long-term survival is difficult to discern, especially due to the presence of additional factors such as body weight changes over time or dose titrations. Those results were reflected in the US drug label issued on April 27, 2015 within the Pediatric Use section [6]. Furthermore FDA stated the effect of REVATIO® on the risk of death with long-term use in adults is unknown and has required Pfizer to evaluate REVATIO®'s effect on the risk of death in adults with PAH [7]. A clinical trial entitled "Effects of Oral Sildenafil on Mortality in Adults With PAH" (AFFILIATE, NCT02060487) has started in September 2014.

The objective of this modeling and simulation analysis was to assess the interpretation of the AFFILIATE mortality trial in adults and not to challenge the assumption of an exposure–mortality relationship nor its extrapolation to the adult population. This work characterized the exposure–mortality relationship, and utilized it to simulate the AFFILIATE trial to assess its interpretation according to varying conditions. The work was performed in two parts: (1) to adequately characterize the exposure–mortality relationship observed in STARTS-1 and STARTS-2 and (2) to simulate the AFFILIATE mortality trial in adults under both assumptions: absence or presence of an exposure–mortality relationship, utilizing the previously developed model in the latter case. It was particularly important to account for longitudinal covariates changes along time such as body weight, and drug exposure in the analysis of STARTS-1 and STARTS-2 survival data as previously described [8]. It was also important to account for potential alterations that may impact the AFFILIATE trial conduct.

Methods

Trials and data

Data were obtained from a 16-week, placebo-controlled, dose-ranging study evaluating the effects of oral sildenafil in treatment-naive children, aged 1–17 years, with PAH (STARTS-1) and from its extension study (STARTS-2) [1]. According to Table 1, in STARTS-1, patients $\leq$ 20 kg were randomized 1:1:2 to placebo, medium-, or high-dose sildenafil, respectively. Patients > 20 kg were randomized 1:1:1:1 to placebo and sildenafil low-, medium-, or high-dose. Only patients who completed STARTS-1 were eligible to enroll in the extension study (STARTS-2). Patients who received sildenafil in STARTS-1 were maintained on the same sildenafil dose; whereas placebo treated patients were re-randomized to receive low-, medium-, or high-dose sildenafil monotherapy. Nominal doses were ranging from 10 mg to 80 mg tid and adjusted based on patients' body weight along STARTS-2 conduct. Up and down titrations were also allowed at the physician discretion without exceeding the high-dose, and the lowest dose being 10 mg. Available survival data were used as of August 4, 2011 when the data monitoring committee recommended discontinuation of the 40 mg and 80 mg doses, as well as the 20 mg dose in children with body weight < 20 kg. This brought all pediatric patients to the already recommended dose regimen approved in the EU [9]. One-third of patients had IPAH/HPAH, the two other thirds were composed of patients with either surgical repair, congenital heart disease (shunt) or D-transposition of the great arteries and grouped as CHD PAH. Full dosing history, death, and censoring dates were extracted from the database, as well as covariates to be tested as prognostic factors.

Survival model

The survival model aimed at relating prognostic factors (baseline or time varying covariates) to the clinical endpoint (survival time). The hazard function h(t) is dependent on a set of p covariates which may be time-dependent $x_1(t)$, $x_2(t)$, $x_3(t)$,…, $x_p(t)$ with the size of the effect measured by a set of estimated parameters β_1, β_2, β_3,…, β_p and $\exp(h_0(t))$ characterizing the estimated baseline hazard as expressed:

$$h(t) = exp\left[h_0(t) + \sum_{j=1}^{p} \beta_j \cdot x_j(t)\right]. \quad (1)$$

The probability of survival up to time t, S(t) is given by:

$$S(t) = exp\left[\int_0^t h(u) \cdot du\right]. \quad (2)$$

The survival model was a "fit for purpose" model with two covariate effects. Indeed sildenafil exposure represented by longitudinal average concentration at steady state (Cav,ss) was included by default into the model to address the issue raised by the FDA [7]. It was derived using mean clearance values from a previously developed population PK model and individual time-varying values of dose and body weight [10]. Although PK samples and thus individual exposure (derived from empirical Bayesian estimates) were available for STARTS-1, only typical predictions could reasonably be used for survival modeling because there was no PK samples during the long-term extension (STARTS-2) when death events occurred, leading us to assume that the individual exposure measured in

Table 1 Sildenafil thrice daily dose in STARTS-1 and STARTS-2

Body weight (kg)	Low dose (mg)	Medium dose (mg)	High dose (mg)
≥ 8–20	NA	10	20
> 20–45	10	20	40
> 45	10	40	80

NA not available, no low dose group for the ≥ 8–20 kg body weight category

Patients ≤ 20 kg were randomized 1:2:1 to placebo, medium-, and high-dose sildenafil, respectively. Patients > 20 kg were randomized 1:1:1:1 to placebo and sildenafil low-, medium-, and high-dose groups

Actual doses administered within a dose group were dependent on body weight

Patients completing STARTS-1 were eligible to enroll in an extension study STARTS-2. Patients who received sildenafil monotherapy in STARTS-1 were maintained on the same sildenafil dose they received while in the 16-week study. Placebo treated

patients were randomized to receive low-, medium-, or high-dose sildenafil monotherapy

STARTS-1 was not necessarily a precise reflection of the unobserved exposure in STARTS-2. Cav,ss values in this analysis ranged from 4 to 335 ng/mL. The effect on Cav,ss on survival was analyzed in various ways. Cav,ss was set to 0 in placebo subjects in STARTS-1. First, as continuous variable, linear and non-linear models were considered. Second, as categorical variable according to a threshold estimated during the model development, no effect below the threshold and maximum effect above the threshold. The threshold was estimated using log-likelihood profiling. Finally, as the combination of categorical and continuous variable: no effect below the threshold and concentration-dependent effect above the threshold. An exploratory data analysis showed that etiology had a big impact on survival and was thus the second covariate included in the model as: IPAH/HPAH versus CHD [11]. The survival model was developed using the non-linear mixed effects modeling program (NONMEM) (Version 7.2.0; GloboMax, Hanover, MD).

The survival model was evaluated using Visual Predictive Checks (VPC). Survival times were simulated 2000 times for the same number of patients as in the dataset used for model development; Cav,ss and body weight histories, and baseline covariates were sampled from the original dataset. The mean parameter values were used (uncertainty in parameter estimates not accounted for). Observed survival distributions (Kaplan–Meier curve) were compared with the predictive distributions by the model (90% prediction interval) in different sub-populations of interest. Posterior Predictive Checks (PPC) of the HR high-dose versus low-dose and the HR according to relevant Cav,ss cut-off were also performed using 2000 replicates [12]. Uncertainty in parameter estimates was accounted for: parameter values for the survival model were sampled from the estimated mean values and variance–covariance matrix. For each simulated replicate, the HR of e.g. high-dose versus low-dose was estimated by fitting a Cox proportional hazard regression model. The 90% prediction interval of the HR was derived and compared to the observed one. The model evaluation was performed in R version 3.0.1.

Clinical trial simulations

Outcomes of virtual survival trials in an adult PAH population treated with sildenafil were simulated using the study design elements (AFFILIATE) and the survival model developed on pediatric data. The base design consisted of three dosing arms 5, 20 and 80 mg given three times a day, the assumptions used for sample-size calculation were: a recruitment rate of 100 subjects per year, a 2 years mortality rate of 15% (0.0002/day) in all subjects and a 20% dropout-rate. The baseline body weight distribution was: 72.0 ± 17.5 kg (minimum = 38 kg, maximum = 137 kg) based on previous sildenafil trials performed in adults [13, 14]. No body weight changes over time were assumed in the simulated adult PAH population. According to the base design (scenario 1), 131 events were needed to show non-inferiority of 80 mg versus 5 mg with 90% power requiring 429 subjects (143 per arm); under those conditions the trial would last 7.7 years. This predicted number of 131 events assumed that the AFFILIATE trial would be conducted without interim analysis; all simulations were performed under the same assumption. In all tested scenarios, the number of events (131) and the total number of subjects (429) were kept constant. Two sets of simulation scenarios were considered:

1. A "No exposure–mortality" set in which no exposure–mortality relationship was assumed, e.g. 5 mg, 20 mg and 80 mg incur the same hazard, with scenario 1 as base case
2. An "Exposure–mortality" set where exposure–mortality relationship was adopted from the survival model developed on pediatric data, 80 mg having a worse outcome compared to 5 mg, with scenario 8 as base case assuming an homogenous population of IPAH/

HPAH patients. As the lowest tested dose was 10 mg in STARTS-1 and STARTS-2, the same exposure–mortality response was assumed between 5 mg and 10 mg.

In each setting, a varying set of assumptions was implemented to reflect the impact confounding factors could have on the outcome of the primary endpoint i.e. overall survival:

- Disease progression assuming linear increments of the hazard over time of 0.000068/day (equivalent to a 2 years mortality rate of 5%) each year corresponding to a linear increase of hazard of $1.88.10^{-7}$/day each day, in scenarios 2 and 10.
- Add-on therapy that over-rules the underlying mortality hazard ranging from 2 years mortality rate of 5% (improving add-on therapy) to 30% (worsening add-on therapy) applied randomly to 50% of patients after 2 years of treatment in scenarios 3 and 11.
- A doubled dropout-rate of 40% in scenarios 4 and 13.
- To account for the impact of etiology, it was assumed 33% randomly selected subjects were CHD-PAH subjects and 67% were IPAH/HPAH, a HR of IPAH/HPAH versus CHD of 6.5 was estimated on the pediatric dataset, this was reflected in scenarios 5a and 14a. PAH etiology in adults is different compared to pediatric patients with 25% of adult PAH patients presenting PAH associated connective tissue disease (CTD) [11]. Survival information on this etiology was obtained from the REVEAL registry and a HR of 1.59 of CTD versus IPAH/HPAH subjects was used in the simulations [11]. Scenarios 5b and 14b were based on 33% randomly selected subjects presenting CTD etiology while 67% were IPAH/HPAH.
- Background therapy with bosentan applied to 50% randomly selected subjects: sildenafil exposure was assumed to be reduced by 67% due to the drug-interaction with bosentan [15].

Combination scenarios (6, 7, 15, 16, 17) started with one assumption and switched 50% of subjects to another assumption after 2 years of treatment. In the case of scenario 6, disease progression according to scenario 2 was assumed in all subjects, and after 2 years of treatment 50% of subjects remained under the same assumption while the remaining 50% were switched to an improving add-on therapy as assumed in scenario 3a. For combination scenario 17 (add-on therapy and bosentan), confounding factors were independently randomly allocated to 50% of patients leading to four groups: no confounding, bosentan only, add-on therapy only and both confounding factors.

For scenarios 2, 5, 6, 7, 9, 10, 15 and 16, the default hazard was obtained from alternative models: that do not account for any covariate effect (2, 6, 7), that only account for etiology effect (5) or that only account for Cav,ss effect (9, 10, 15, 16).

500 replicates of each scenario were run. Similarly to the PPC, for each replicate, the HR of 80 mg versus 5 mg was estimated with its 95% confidence interval (CI) by fitting a Cox proportional hazard regression model. Respective powers for non-inferiority and superiority were defined based on two decision criteria following the planned statistical analysis of the AFFILIATE trial:

- Non-inferiority: percentage of replicates with 95% CI of HR upper boundary < 2.
- Superiority: percentage of trials with 95% CI of HR lower boundary > 1.

All simulations were performed and post-processed in R version 3.0.1.

Results

Survival model

A total of 234 patients were randomized in STARTS-1, of which 220 patients continued into the long-term extension STARTS-2 [5]. The survival dataset which combines STARTS-1 and STARTS-2 data comprised 232 patients, indeed covariate information was missing for two subjects. The dataset contained 37 death events consistent with the survival data reported in the drug label: 5 in the low-dose group (9%), 10 in the medium-dose group (14%) and 22 in the high-dose group (22%) [6]. Up-titrations (not due to body weight changes) occurred in 52 subjects, most of them (n = 28) in the low-dose group while there were only 6 down-titrations [5]. During model development, initial visual predictive checks (VPCs) showed a consistent under-prediction of the survival during the first year. Indeed the earliest mortality event occurred after 324 days of treatment. Therefore the hazard was linearly interpolated as described by Eq. (3) between time 0, when the hazard was null h(0) = 0, and an estimated onset time (ONSET) where the hazard took the value h(ONSET) as derived by Eq. (4) below. Survival in pediatrics was found to be mainly impacted by etiology with idiopathic or heritable PAH (IPAH/HPAH) patients having a worse outcome ($p = 7.10^{-8}$ when removing etiology effect from final model) and to a minor extent by drug exposure (p = 0.040 when removing exposure effect from final model). The effect of exposure was best described as the combination of categorical and continuous variable: no effect below a certain Cav,ss threshold and concentration-dependent effect above the threshold. The optimal Cav,ss threshold was determined by log-likelihood profiling at 89 ng/mL

(Fig. S1). In each subject i, the hazard at any time t was expressed as follows:

$$h(t<ONSET)_i = \frac{t}{ONSET}\cdot\exp\left(h_0 + Etiology_i\cdot\beta_{ETIO} + Cav,ss(t)_i\cdot\beta_{CONC}\right) \quad (3)$$

$$h(t\geq ONSET)_i = \exp\left(h_0 + Etiology_i\cdot\beta_{ETIO} + Cav,ss(t)_i\cdot\beta_{CONC}\right) \quad (4)$$

h_0 is the parameter that describes the baseline hazard; $Etiology_i$ = 1 in IPAH/HPAH subjects and 0 in other ones; β_{ETIO} gives the magnitude of etiology effect in IPAH/HPAH subjects; Cav,ss designates the typical sildenafil average concentration at steady state; β_{CONC} gives the magnitude of the drug effect by ng/mL above a threshold of 89 ng/mL; the drug effect is null below that threshold. As typical Cav,ss only depends on dose, body weight and drug clearance, it is possible to derive it for a given dose and to position it with respect to the threshold of 89 ng/mL: values for subjects treated with 40 or 80 mg tid would exceed 89 ng/mL while values for subjects treated with 10 mg tid would be below that threshold, values for subjects < 20 kg treated with 20 mg tid would also exceed 89 ng/mL, values for subjects ≥ 20 kg treated with 20 mg tid could belong to both categories depending on their body weight and drug clearance values. In terms of pharmacodynamics, 89 ng/mL is a Cav,ss level at which maximum hemodynamic (pulmonary vascular resistance) and exercise capacity (VO2 peak) are nearly achieved [16, 17].

Parameters (h_0, ONSET, β_{ETIO}) were well estimated with relative standard errors (RSE) less than or close to 20% except for β_{CONC} (RSE of 43%) (Table 2). The HR of IPAH/HPAH subjects versus subjects with congenital heart defect (CHD) (assuming same Cav,ss value) was 6.5 while the HR between concentrations below the threshold Cav,ss and the highest Cav,ss value (335 ng/mL) was 3.8 (assuming same etiology). As illustrated by the VPCs (Figs. S2, S3 and S4), the final model had good predictive performance in simulating the survival distribution (Kaplan–Meier plots) in the different treatment, concentration and etiology groups. A posterior predictive check (PPC) of the HR of high-dose versus low-dose resulted in a 90% prediction interval of 0.7–4.4, which included the observed value of 3.4 [5]. Similarly, the PPC of the HR of average exposure $>$ 89 ng/mL versus average exposure $\leq$ 89 ng/mL gave a 90% prediction interval of [0.8; 3.2] that also included the corresponding observed value of 1.5.

Table 2 Parameters of the survival model

Parameters	Unit	Estimate	SE	RSE (%)
h_0	1	− 10.5	0.361	3.4
ONSET	days	471	1.09	0.2
β_{ETIO}	1	1.87	0.383	20.5
Threshold Cav,ss*	ng/mL	89	–	–
β_{CONC}	mL/ng	0.00401	0.00171	42.6

*Value determined by log-likelihood profiling

h_0: base hazard at Cav,ss = 0 in CHD-PAH subjects

ONSET: estimated onset time

β_{ETIO}: effect in IPAH/HPAH subjects

β_{CONC}: slope of the Cav,ss dependent effect

Clinical trial simulations

Figure 1 illustrates the evolution of the hazard over time according to the survival model in different virtual adult subjects. It shows the magnitude of both respective effects, with etiology having the biggest impact. Figures 2 and 3 show how confounding factors, namely disease progression, add-on therapy (including bosentan), dropout-rate, or etiology, would impact the hazard over time in a typical virtual subject for both sets of simulations: "No exposure–mortality" or "Exposure–mortality", respectively. The base case scenario (scenario 1 obtained with the base design: three dosing arms 5, 20 and 80 mg, an recruitment rate of 100 subjects per year, a 2 years mortality rate of 15% in all subjects and a 20% dropout-rate) and scenarios derived from it (3a–3e) assumed a constant hazard over time from time = 0 (horizontal lines in Fig. 2). Other

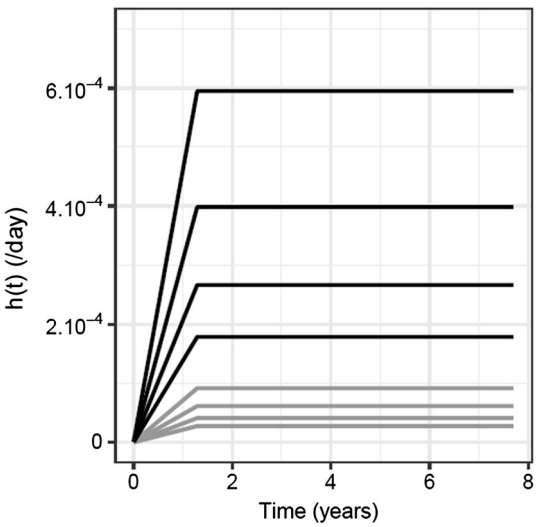

Fig. 1 Hazard over time at different Cav,ss values and for both etiologies. Grey lines represent the hazard-time profiles at Cav,ss values of 5 (bottom grey line), 100, 200 and 300 (top grey line) ng/mL respectively in CHD subjects. Black lines represent the hazard-time profiles at Cav,ss values of 5 (bottom black line), 100, 200 and 300 (top black line) ng/mL in IPAH/HPAH subjects

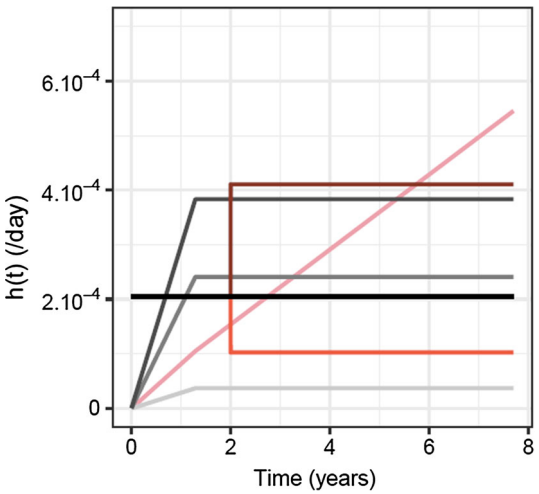

Fig. 2 Impact of confounding factors on the time-varying hazard in "No exposure–mortality" scenarios. The black line represents the hazard according to the base case scenario (scenario 1). The pink line represents how disease progression (scenario 2), would impact the hazard over time. The red and brown lines represent how improving add-on therapy (scenario 3b) and worsening add-on therapy (scenario 3e), respectively, would impact the hazard over time after 2 years of treatment. Grey lines account for the etiology difference: dark grey in CTD subjects, grey in IPAH/HPAH subjects and light grey in CHD subjects based on the survival model (Color figure online)

scenarios which are based on the developed survival model (or its alternative models) assumed a linear interpolation between time 0, when the hazard was null h(0) = 0, and the estimated onset time (ONSET = 471 days) when the hazard took the value h(471) as derived by Eqs. (3) and (4).

Table 3 shows the simulation results in all evaluated "No exposure–mortality" scenarios. Non-inferiority would be shown in 90% of trial replicates even in presence of confounding factors and combination of confounding factors. The survival trial can be considered as robust under those conditions although the length of the trial would be impacted extending likely from 7 to almost 10 years to reach the required 131 events.

Figure 3 shows how confounding factors would impact the hazard over time in typical virtual subjects receiving either 5 mg or 80 mg tid in the "Exposure–mortality" set of simulations. Table 4 shows the simulation results in all "Exposure–mortality" scenarios (80 mg has a worse outcome compared to lower doses). If the exposure–mortality relationship observed in children would apply to adults (scenario 8), the risk of wrongly concluding non-inferiority of 80 mg versus 5 mg is predicted to be low (8% of replicates). Scenario 9 represents the unlikely scenario where etiology differences would not be accounted for but gives the upper boundary of the clinical trial length: 13.3 years. Superiority (80 mg worse than 5 mg) would still be predicted in 80–90% of simulated trials even in presence of confounding factors such as etiology: 33% of the population being composed of either CTD subjects in scenario 14b or other CHD subjects in scenario 14a; or an increased dropout rate (scenario 13). Similarly to the "No

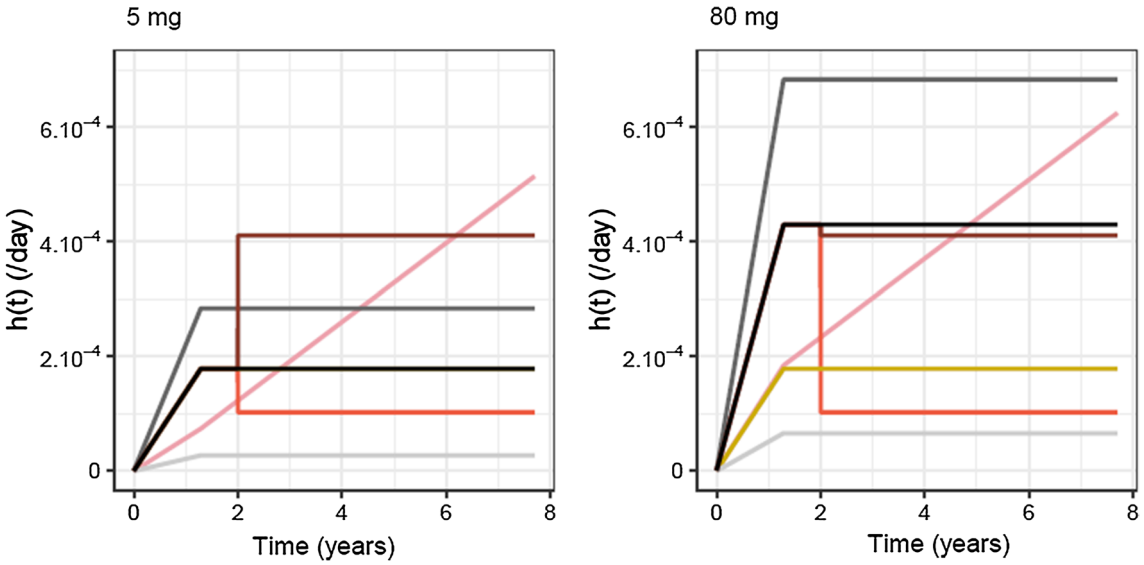

Fig. 3 Impact of confounding factors on the time-varying hazard in "Exposure–mortality" scenarios. Black lines represent the hazard at 5 mg or 80 mg in a typical IPAH/HPAH subject with a body weight of 72 kg (scenario 8). The pink and olive lines represent how disease progression (scenario 10), and combination with bosentan (scenario 13), respectively, would impact the hazard over time. At 5 mg, the olive line is confounded with the black line (the value of Cav,ss does not matter as long as it is below the threshold of 89 ng/mL). Red and brown lines represent how improving add-on therapy (scenario 11b) and worsening add-on therapy (scenario 11f), respectively, would impact the hazard over time after 2 years of treatment. Grey lines account for the etiology difference: dark grey in CTD subjects and light grey in CHD subjects (Color figure online)

Table 3 Simulations results for "No exposure–mortality" scenarios

Scenario	Mortality trial design features-alterations					Simulation results		
	Default hazard (%/2 years)*	Impact of	Hazard change (%/2 years) and details	Population affected by change (%)	Beginning of change	Duration (years)	HR 80 versus 5 mg Median [90% prediction interval]	% trials concluding non-inferiority
1, base case	15.0%					7.7	0.99 [0.71;1.41]	90.6
2	7.7%	Disease progression	prop. with time (+5% each year)	100%	At start	7.7	0.97 [0.71;1.49]	87.4
3a	15.0%	Add-on	switch to 5%	50%	At 2 years	9.5	0.99 [0.69;1.39]	91.0
3b	15.0%	Add-on	switch to 7.5%	50%	At 2 years	8.8	0.99 [0.69;1.47]	88.6
3c	15.0%	Add-on	switch to 10%	50%	At 2 years	8.4	1.00 [0.70;1.41]	90.8
3d	15.0%	Add-on	switch to 22.5%	50%	At 2 years	7.1	1.00 [0.72;1.44]	89.0
3e	15.0%	Add-on	switch to 30%	50%	At 2 years	6.6	1.00 [0.72;1.41]	91.4
4	15.0%	Dropout	doubled	100%	At start	8.5	1.00 [0.72;1.42]	90.0
5a	17.6%	Eiology	HR IPAH/HPAH versus CHD 6.5	33% CHD	At start	9.9	0.99 [0.71;1.43]	88.2
5b	17.6%	Etiology	HR CTD versus IPAH/HPAH 1.59	33% CTD	At start	6.9	1.02 [0.70;1.42]	88.6
		Combination scenarios						
6: 2+3a	7.7%	Disease prog./add-on	prop. with time (+ 5% each year)/ switch to 5%	100%/50%	At start/at 2 years	8.2	1.01 [0.71;1.45]	89.2
7: 2+3e	7.7%	Disease prog./add-on	prop. with time (+ 5% each year)/ switch to 30%	100%/50%	At start/at 2 years	6.9	0.97 [0.69;1.42]	92.0

*Value in IPAH/HPAH subjects for scenarios 5a and 5b

Scenario 2, 6 and 7 were performed with an alternative survival model neither accounting for etiology nor for Cav,ss effects

Scenarios 5a and 5b were performed with an alternative model without accounting for Cav,ss effects

Table 4 Simulations results for "Exposure–mortality" scenarios

Scenario	Mortality trial design features-alterations					Simulation results			
	Default hazard (%/2 years)* At Cav,ss = 0	Impact of	Hazard change (%/2 years) and details	Population affected by change (%)	Beginning of change	Duration (years)	HR 80 versus 5 mg Median [90% prediction interval]	% trials concluding	
								NI	SUP
8, Base case	13.1%	Exposure–response (exp-resp) only, 100% IPAH/HPAH		100%	At start	7.4	2.47 [1.25;5.30]	8.0	85.2
9	5.4%	Exp-resp only	Etiology not accounted for in model	100%	At start	13.3	2.58 [1.24;5.39]	6.2	88.8
10	5.4%	Disease progression	Prop. with time (+5% each year)	100%	At start	7.7	1.58 [1.00;2.75]	25.4	56.2
11a	13.1%	Add-on	Switch to 5%	50%	At 2 years	9.3	2.03 [1.12;4.01]	12.8	79.0
11b	13.1%	Add-on	Switch to 7.5%	50%	At 2 years	8.7	2.03 [1.09;3.78]	12.8	78.0
11c	13.1%	Add-on	Switch to 10%	50%	At 2 years	8.3	1.99 [1.10;3.62]	15.4	76.6
11d	13.1%	Add-on	Switch to 15%	50%	At 2 years	7.6	1.85 [1.05;3.59]	14.8	71.6
11e	13.1%	Add-on	Switch to 22.5%	50%	At 2 years	7.0	1.86 [1.08;3.22]	16.2	72.0
11f	13.1%	Add-on	Switch to 30%	50%	At 2 years	6.7	1.83 [1.03;3.18]	17.6	71.2
12	13.1%	Bosentan	Exposure reduced by 67%	50%	At start	8.2	1.65 [0.97;2.92]	25.0	62.4
13	13.1%	Dropout	Doubled	100%	At start	8.1	2.41 [1.21;5.06]	8.4	85.8
14a	13.1%	Etiology	HR IPAH/HPAH versus CHD 6.5	33% CHD	At start	9.6	2.30 [1.19;4.55]	8.8	82.8
14b	13.1%	Etiology	HR CTD versus IPAH/HPAH 1.59	33% CTD	At start	6.7	2.44 [1.19;5.25]	9.0	86.2
		Combination scenarios							
15:10+11a	5.4%	Disease prog./add-on	Prop. with time (+5% each year)/ switch to 5%	100%/50%	At start/at 2 years	8.1	1.38 [0.92;2.14]	42.2	37.4
16: 10+11f	5.4%	Disease prog./add-on	Prop. with time (+5% each year)/ switch to 30%	100%/50%	At start/at 2 years	6.9	1.37 [0.89;2.13]	44.2	32.0
17: 11b+12	13.1%	Bosentan/add-on	Exposure reduced by 67%/switch to 7.5%	50%/50%	At start/at 2 years	9.6	1.45 [0.93;2.22]	36.6	42.8

Exposure–mortality relationship adopted from the pediatric survival model

NI non-inferiority, *SUP* superiority

*Value in IPAH/HPAH subjects for scenarios 8 and 11, 12, 13, 14, 17

Scenarios 8, 11, 12, 13, 17 assumed 100% subjects were IPAH/HPAH

Scenarios 9 and 10, 15 and 16 were performed with an alternative model without accounting for the etiology

exposure–mortality" scenarios, the length of the trial would vary from 7 to 10 years. However in a number of cases the ability of showing a potential detrimental effect of 80 mg versus 5 mg would drop. In particular in the scenarios which assume disease progression (scenario 10), improving or worsening add-on treatment that over-rules the mortality effect related to sildenafil (scenarios 11), combination therapy with bosentan (scenario 12) the survival trial outcome may be misinterpreted. Figure S5 shows how add-on therapy and its associated hazard would impact the ability to show non-inferiority or superiority. In the "No Exposure–mortality" scenarios, add-on therapy would not have a large impact on the results, as non-inferiority would still be concluded in 90% of simulated trials. However, for the "Exposure–mortality" scenarios, the trial ability would decrease from 85 to 71% under an increasing hazard due to add-on therapy. The trial length to reach 131 events would similarly be impacted in both sets of scenarios. Scenarios 15, 16 and 17 show a higher risk of misinterpretation in the case of combination of confounding factors: add-on therapy combined with either disease progression or combination therapy with bosentan.

Discussion

This work describes the use of modeling and simulation to assess the interpretation of a mortality trial by testing the impact of the presence of confounding factors that could arise during the trial conduct. This work was based on the FDA interpretation that the exposure–mortality relationship observed in pediatric patients could be drug-related and thus could also apply to adult patients. Results from a model-based assessment of survival in pediatric PAH patients allowed to construct a simulation framework for adult PAH patients, according to FDA's rationale. The adequate characterization of survival drivers using a longitudinal analysis made it possible to assess the outcome of a survival trial in adults and the impact external confounding factors can have on interpretation.

Survival in pediatric patients was found to be mainly impacted by etiology with idiopathic or heritable PAH (IPAH/HPAH) patients having a worse outcome as illustrated by Figures S2 and S3 and to a minor extent by drug exposure. This means that once we account for etiology, the impact of drug exposure is much less pronounced. The linear interpolation assumed between the beginning of the STARTS-1 trial and the first event as described by Eq. (3) may be questionable. But applying a high hazard from the beginning, as represented by Eq. (4), would predict mortality events when none occurred up to 324 days, while all the deaths happened in the subsequent 6.5 years. The linear interpolation appeared to be a conservative approach by assuming that the hazard would start increasing concomitantly with treatment start in accordance with FDA's rationale, i.e. mortality potentially induced by sildenafil, and allowed a good description of the observed data by the model. The predictive performance of the model was assessed by visual and posterior predictive checks and allowed the use of the model to simulate the AFFILIATE trial.

Tested confounding factors were: etiology, disease progression, dropout-rate, add-on therapy and combination therapy with bosentan. Disease progression and dropout-rate are both relevant confounding factors for a survival trial that likely lasts more than 7 years. PAH treatment often consists of dual or triple combinations between endothelin receptor antagonists (e.g. ambrisentan, bosentan, macitentan), phosphodiesterase type-5 inhibitors (e.g. sildenafil, tadalafil) and prostacyclins (e.g. epoprostenol, treprostinil), therefore it was pertinent to evaluate the potential impact of add-on therapies [11]. For etiology, patients with CTD represent 25% of adult PAH patients and are known to have a worse outcome than IPAH/HPAH patients, a HR of 1.59 was used; [11] other PAH etiologies may have a worse outcome such as PAH associated with portal hypertension [4, 11] or a better outcome as observed in CHD patients [5]. In subjects with PAH other than IPAH/HPAH or PAH associated with CTD, it was decided to leverage the information obtained from the survival model (HR IPAH/HPAH versus CHD of 6.5).

This work, as the AFFILIATE trial, is substantiated by the assumptions made notably on the unfavorable dose-mortality relationship as observed in STARTS-2.

Within the "No exposure–mortality" set (assuming no treatment difference), confounding factors are less likely to have an impact; while in the "Exposure–mortality" set, confounding factors would likely impact interpretation, especially since confounding factors often occur simultaneously. Among confounding factors, disease progression and add-on therapy (bonsentan or another treatment) would more likely impact interpretation. For instance, adding a treatment which would improve survival after 2 years (scenario 11a) would make sense and the ability of showing an unfavorable exposure–mortality relationship would still be acceptable (79.0%) but it would considerably drop (37.4%) in case of a combination of confounding factors: improving add-on therapy/disease progression (scenario 15).

The case of bosentan is particularly clinically relevant, as the combination of bosentan and sildenafil has been reported to be beneficial in terms of survival [4]. From a pharmacokinetic perspective, the CYP3A4 interaction between sildenafil and bosentan has been previously reported and revealed a two-third reduction of sildenafil exposure in the presence of bosentan [18]. In case of an

adult sildenafil survival trial with bosentan as background therapy, the trial ability to conclude a potential detrimental effect of 80 mg versus 5 mg of sildenafil would only be 62% assuming a 67% reduction in sildenafil exposure for subjects receiving a dual therapy sildenafil/bosentan. Indeed, a reduced sildenafil exposure at 80 mg due to the combination with bosentan would result in a reduced risk compared with sildenafil monotherapy at 80 mg. Hence, a median HR of 1.65 versus 2.47 would translate into a lower event rate and a longer trial duration under combination therapy. Study COMPASS-2 (NCT00303459) investigated their combination with time to first morbidity/mortality event as the primary outcome [19]. A HR of 0.83 resulted for bosentan in combination with sildenafil compared to sildenafil alone, but the difference was not statistically significant (CI 0.58–1.19).

In both sets of simulations the length of the trial to reach the required number of events (131) would vary considerably from 6.7 to 13.3 years depending on the scenarios and confounding factors having an impact, even though the recruitment rate of 100 subjects per year is maintained along the trial conduct. For instance, switching to an add-on therapy with an associated lower hazard than the one from the base case would slow down the occurrence of events and thus increase the trial duration while the exact opposite effect would be expected if the add-on therapy was associated with a higher hazard than the base case hazard. Recruiting patients with CTD (worse prognostic factor) would lead to a faster occurrence of events and thus reduce the trial duration while recruiting CHD patients (better prognostic factor) would increase the trial duration compared to the base case scenario.

Besides, increasing the number of events by prolonging the trial would not or only marginally improve the trial ability in the most compromised scenarios (data not shown here).

This work raises the importance of carefully controlling confounding factors to assess the risk of potentially wrongly concluding non-inferiority when designing such a trial. In the case of the AFFILIATE trial the stratification by treatment at study entry (naïve versus on treatment) and by etiology (IPAH/HPAH versus CHD) will reduce part of this risk but will not control confounding factors occurring during the trial conduct.

Alongside assessing the trial design, it has to be noted that in this particular case pediatric data informed the adult indication while in most cases in clinical drug development it is the reverse.

This work contributes to the design of an adult mortality trial and the assessment of its interpretation. A similar methodology could be applied to other indications where confounding factors are likely having an impact on the conduct and readout of a clinical trial.

Conclusions

A clinical trial entitled “Effects of Oral Sildenafil on Mortality in Adults With PAH” (AFFILIATE) is ongoing. Clinical trial simulations were performed to assess the design of the survival trial in adults (AFFILIATE, expected to last approximately 8 years) according to both assumptions: absence or presence of an exposure–mortality relationship and to quantify the impact of confounding factors on its readout. Simulations showed that the trial would be robust in most conditions. But its interpretation will depend on the extent of confounding factors such as additional treatment attempt to control disease progression.

Acknowledgements This work was sponsored by Pfizer. Mathilde Marchand, an employee of Integrated Drug Development, Certara, Princeton, NJ, USA was paid consultant to Pfizer and performed the QC for this manuscript.

Author contributions PC and LH wrote the manuscript, PC, XG, RB, LC and LH designed the work, PC and LC analyzed the data.

Compliance with ethical standards

Conflict of interest PC, RB and LC are employees of Genentech/Roche and were paid consultants of Certara/Pharsight Consulting Services to Pfizer in connection with the analyses and development of this manuscript. XG is an employee of Alexion Pharmaceuticals. LH is an employee of Pfizer.

Data sharing statement Upon request, and subject to certain criteria, conditions and exceptions (see https://www.pfizer.com/science/clinical-trials/trial-data-and-results for more information), Pfizer will provide access to individual de-identified participant data from Pfizer-sponsored global interventional clinical studies conducted for medicines, vaccines and medical devices (1) for indications that have been approved in the US and/or EU or (2) in programs that have been terminated (i.e., development for all indications has been discontinued). Pfizer will also consider requests for the protocol, data dictionary, and statistical analysis plan. Data may be requested from Pfizer trials 24 months after study completion. The de-identified participant data will be made available to researchers whose proposals meet the research criteria and other conditions, and for which an exception does not apply, via a secure portal. To gain access, data requestors must enter into a data access agreement with Pfizer.

References

1. Barst RJ, Ivy D, Gaitan G, Szatmari A, Rudzinski A (2012) A randomized, double-blind, placebo-controlled, dose-ranging study of oral sildenafil citrate in treatment-naive children with pulmonary arterial hypertension. Circulation 125:324–334
2. Savarese G, Paolillo S, Costanzo P, D’Amore C, Cecere M (2012) Do changes of 6-minute walk distance predict clinical events in patients with pulmonary arterial hypertension? JACC 60:1192–1201

3. Galie N, Manes A, Negro L, Palazzini M, Bacchi-Reggiani ML (2009) A meta-analysis of randomized controlled trials in pulmonary arterial hypertension. Eur Heart J 30:394–403
4. Haworth SG, Hislop AA (2009) Treatment and survival in children with pulmonary arterial hypertension: the UK Pulmonary Hypertension Service for Children 2001–2006. Heart 95:312–317
5. Barst RJ, Beghetti M, Pulido T, Layton G, Konourina I, Zhang M, Ivy DD (2014) STARTS-2 long-term survival with oral sildenafil monotherapy in treatment-naïve pediatric pulmonary arterial hypertension. Circulation 129:1914–1923
6. FDA (2017) sildenafil REVATIO® drug label. 7-31-2017. https://www.accessdata.fda.gov/drugsatfda_docs/label/2017/203109s011,022473s011,021845s020lbl.pdf. Accessed 4 Dec 2018
7. FDA (2012) Drug Safety Communication. 8-30-2012. http://www.fda.gov/Drugs/DrugSafety/ucm317123.htm. Accessed 5 Dec 2018
8. Vu TC, Nutt JG, Holford NHG (2012) Disease progress and response to treatment as predictors of survival, disability, cognitive impairment and depression in Parkinson's disease. Br J Clin Pharmacol 74:284–295
9. FDA (2012) Clinical Review. 11-30-2012. http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/DevelopmentResources/UCM320472.pdf. Accessed 5 Dec 2018
10. Watt S, Hayashi N, Harnisch L, Gao X (2010) Population pharmacokinetics of sildenafil in paediatric and adult patients with pulmonary arterial hypertension. Eurpoean society of cardiology. http://spo.escardio.org/eslides/view.aspx?eevtid=40&fp=P4506. Accessed 5 Dec 2018
11. Mc Goon MD, Miller DP (2012) REVEAL: a contemporary US pulmonary arterial hypertension registry. Eur Respir Rev 123:8–18
12. Yano Y, Beal SL, Sheiner LB (2001) Evaluating pharmacokinetic/pharmacodynamic models using the posterior predictive check. J Pharmacokinet Pharmacodyn 20:171–192
13. Galiè N, Ghofrani HA, Torbicki A, Barst RJ, Rubin LJ, Badesch D, Fleming T, Parpia Burgess G, Branzi A, Grimminger F, Kurzyna Simonneau G for the Sildenafil Use in Pulmonary Arterial Hypertension (SUPER) Study Group (2005) Sildenafil citrate therapy for pulmonary arterial hypertension. N Engl J Med 353:2148–2157
14. Simonneau G, Rubin LJ, Galiè N, Barst RJ, Fleming TR, Frost AE, Engel PJ, Kramer MR, Burgess G, Collings L, Cossons N, Sitbon O, Badesch DB, PACES Study Group (2008) Addition of sildenafil to long-term intravenous epoprostenol therapy in patients with pulmonary arterial hypertension. Ann Intern Med 149:521–530
15. Humbert M (2007) Dual endothelin receptor antagonism: setting standards in PAH. Eur Respir Rev 16:13–18
16. Harnisch L, Chanu P, Gao X, Smith M, Bruno R (2011) Hemodynamics as an additional measure to prove efficacy and to provide a dose rationale for sildenafil in pediatric patients with pulmonary arterial hypertension. Am J Respir Crit Care Med 183(2011):A6281. https://doi.org/10.1164/ajrccm-conference.2011.183.1_MeetingAbstracts.A6281
17. Harnisch L, Hayashi N (2009) Exercise tolerability in children with pulmonary arterial hypertension (PAH): a population PK/PD assessment of the effects of sildenafil. Abstract P3890, European Respiratory Society Annual Congress, Vienna, Austria, Sept 12–16, 2009
18. Paul GA, Gibbs SJR, Boobis AR, Abbas A, Wilkins MR (2005) Bosentan decreases the plasma concentration of sildenafil when coprescribed in pulmonary hypertension. Br J Clin Pharmacol 60:107–112
19. McLaughlin V, Channick RN, Ghofrani HA, Lemarié JC, Naeije R, Packer M, Souza R, Tapson VF, Tolson J, Al Hiti H, Meyer G, Hoeper MM (2015) Bosentan added to sildenafil therapy in patients with pulmonary arterial hypertension. Eur Respir J 46:405–413