Table: Response to TAR-20-171 Reviewer comments

<table>
<tr><th>Comment</th><th>Response</th><th>Location of edits in track-changes revision</th></tr>
<tr><td>Reviewer 1</td><td></td><td></td></tr>
<tr><td>While the authors note that there is no standardized method of examining RWD of pulmonary hypertension, they adopted a very complex algorithm to identify “Group 1 PAH” patients described in methods (pp5-6) and figure 2. In that vein, please address the following:
-Inclusion of the 101 patients that entered in the algorithm on the CTEPH path (Fig 2) and ultimately were determined to be group 1 PAH at the bottom.</td><td>As shown in Figure 2 (p 23), all patients with “Yes” for presence of a CTEPH International Classification of Diseases, Tenth Revision (ICD-10) code were actually excluded from the PAH cohort at the second level of the flowchart. This rule was applied to both pathways proceeding from that second level, so the right side of the schematic should not be thought of as the CTEPH path. On the contrary, the 101 PAH patients at the bottom of the right-hand pathway in this schematic were determined to be PAH as they did not have a CTEPH ICD-10 code (ie, they flow from the “No” decision node at the second level), but they did have a right heart catheterisation (RHC) and also a specific PAH drug prescribed, so there is a high likelihood they were PAH patients.</td><td>–</td></tr>
<tr><td>-Approximately 45% of the cohort was age 60 or greater. It is unusual the group 1 PAH patients to present at that age. What was the age breakdown of the incident patients?</td><td>The ~45% of the cohort ≥60 years of age refers to the age distribution in the overall cohort, and thus does not correspond to the age at which incident patients presented (ie, were first diagnosed with PAH). Regarding the age breakdown of incident patients, mean ± standard deviation (SD) age at first PAH event was 58.0 ± 16.8 years, as reported in the first paragraph of the Patient characteristics section (p 8, lines 9–10). While it’s true that PAH patients presented at a younger age historically, the age at diagnosis has increased over recent decades.[1] In line with our results, the following table from the most recent (2019) UK National Audit of Pulmonary Hypertension reveals that the median age at diagnosis was ≥60 years for PAH subtypes except PAH associated with congenital heart disease or portal hypertension.[2]
Table A 1: Demographics by PAH type,
Adult patients in the longitudinal cohort, Great Britain, 2009-19
<table><tr><th rowspan="2">Diagnostic group</th><th>Patients</th><th colspan="4">Age at diagnosis</th></tr><tr><th>Number</th><th>Lower quartile</th><th>Median</th><th>Upper quartile</th><th>Mean</th></tr><tr><td>IPAH</td><td></td><td></td><td></td><td></td><td></td></tr><tr><td>…without comorbidities</td><td>1,132</td><td>45</td><td>60</td><td>72</td><td>58</td></tr><tr><td>…with comorbidities</td><td>385</td><td>63</td><td>70</td><td>77</td><td>68</td></tr><tr><td>Connective tissue disease APAH</td><td></td><td></td><td></td><td></td><td></td></tr><tr><td>…scleroderma</td><td>1,045</td><td>59</td><td>67</td><td>73</td><td>66</td></tr><tr><td>…excluding scleroderma</td><td>381</td><td>49</td><td>61</td><td>71</td><td>59</td></tr><tr><td>Congenital heart disease APAH</td><td>1,117</td><td>33</td><td>45</td><td>60</td><td>47</td></tr><tr><td>Portal hypertension</td><td>278</td><td>46</td><td>54</td><td>62</td><td>53</td></tr></table></td><td>–</td></tr>
</table>

Comment	Response	Location of edits in track-changes revision
-Page 10: authors comment that they may have missed “some” PAH patients (limes 19-20); however, in 2015 the PH Centers had >3000 PAH patients and the study captured <1800. The “some” would be closer to 40%!	We thank the reviewer for pointing out that our sentence was misleading. Comparison of our study with the PH Audit is complicated because the audit ceased reporting disease-stratified numbers prior to the end of our study period (2017). However, per Table 10 in the 2015 audit there were 6617 PH patients in 2015 in England and Scotland, and 5776 remained after excluding patients at the two centres outside England. Multiplying this number by the 46% of PH patients who had PAH per Table 11 in the audit yields 2656 patients, which is not substantially higher than the 2527 individual patients in our cohort. We have rewritten the paragraph of the Discussion in question to explain this.	Discussion, p 9, lines 11–14
Discussion of hospitalizations (p10, line 53 to p11, line5): Suggest adding some thought to the risk of readmission after 1st (Burger Chest 2014 REVEAL study of hospitalizations) and also the implications of worse survival for PAH related hospitalizations (Burger same study) and all-cause hospitalizations (Benza Chest 2019 REVEAL 2.0).	We agree that these are important points to add, and we have amended the Discussion accordingly. In addition to the references suggested by the reviewer, we have mentioned a retrospective database study supporting the high rate of readmission[3] and a trial-based analysis of the elevated mortality risk following PAH-related hospitalisation.[4]	Discussion, p 10, lines 7–17
Top 20% Analysis: need to add cautionary notes as the data does not distinguish between increased risk (no severity data, no subgroup data such as that of scleroderma-related or POPH, no renal function) versus under treatment (monotherapy vs dual therapy, prostacyclin for high risk).	We concur that this is important context to include, so have added these cautions to the Discussion.	Discussion, p 11, lines 2–8
Lack of data on number and type of PAH medications as well as pharmaceutical cost should be noted as a limitation.	We have now called this out as a specific study limitation in the Discussion.	Discussion, p 10, lines 19–21; p 11, lines 5–8
Reviewer 2		
While the study in general is addressing a pertinent question in the field of PAH, I have a number of concerns regarding the methodology and approach.	We welcome the reviewer’s consideration that our study is relevant and this opportunity to address their concerns.	–

Comment	Response	Location of edits in track-changes revision
Major Concerns:		
1) The selection of the cohort is somewhat confusing. This is made even more confusing by the "cohort selection" section wording (page 4, line 44+), which details selection criteria in the absence of ICD coding, and figures 1 and 2. On the one hand, the wording and figure suggest patients can be either ICD classified, OR have one of the following: RHC, PAH drug, visit to specialty center (Page 4, line 21-33). On the other hand, the "cohort selection" section details multiple criteria to be met to be included in the study in the absence of an ICD-10 code, as detailed by the complex flowchart in Figure 2. By carefully looking through the Figure 2 flowchart, I gather the authors intended to combine ICD coding with additional criteria (RHC, PAH medication, specialty center referral), and to combine those metrics in a different way in the absence of ICD-10 coding, but this needs to be explained in a much clearer fashion, as it forms the basis of this paper's study population. Only in the discussion is it apparent this is how the authors approached defining their cohort. I would also note that, by using Figure 2, a patient could have underlying scleroderma with digital ulcers, or an underlying diagnosis of erectile dysfunction (both common enough in the PAH population), but be excluded from the study based on these diagnoses, so the algorithm used to identify patients in this cohort will inherently miss a significant subset of PAH patients (an issue the authors did, to their credit, discuss in the discussion section).	To help orient the reader we have added a sentence to the first paragraph of the Cohort selection subsection of the Methods explaining that the algorithm combined the various criteria in different ways depending on presence or absence of diagnosis codes. We acknowledge that the algorithm is complex and for this reason we have highlighted in the Methods that this algorithm was crafted by a multidisciplinary team (p 5, lines 5–6); these experts have deep knowledge of the UK healthcare system. We followed best practices to increase the ability of our algorithm to capture all PAH patients during the study period, whether prevalent or incident (though some patients identified as prevalent in the first year of the study period may have been incident, as we have now explained in the Methods), while excluding other forms of PH. There is, unfortunately, no way to guarantee perfect sensitivity and specificity of an algorithm for PAH based on the HES data. However, the correspondence between sex and age distribution of patients in our study with the data reported in the UK PH Audit, which is a mandatory record of care for these patients in the UK, increases our confidence in the performance of the algorithm. It should be acknowledged that no single real-world database is comprehensive, and although the HES database may not be perfect in terms of defining an entirely accurate PAH cohort, it does provide a detailed view of a patient's journey and healthcare resource utilization, which was the main purpose of this study. We have added as a study limitation in the Discussion that we would have failed to capture PAH patients with comorbid conditions that were exclusion criteria.	Methods, p 4, lines 20–21; p 5, lines 1–2 Discussion, p 9, lines 17–19
2) An additional limitation with the algorithm the authors utilize concerns patients with off-label therapy for PAH. For example, a patient with combined WHO-2 and WHO-3 disease could be provided with targeted therapy after a RHC, and be classified as "PAH" by the algorithm. Off-label PAH therapy for non-PAH patients with PH is a routine occurrence, indeed in the US other large studies have indicated that treatment is provided off-label routinely to WHO-2 and WHO-3 group PH patients, even without RHC, at between 40-80% in the US population. (Maron et al. Circulation 2019;139(16):1861-1864) The authors should address this potentially significant limitation in their discussion.	The reviewer raises a very valid observation regarding the off-label prescribing in the US of drugs indicated only for PAH. However, the situation is very different in the UK where the use of PAH-specific medication is well controlled and the National Audit has ensured that essentially 100% of patients receiving these drugs have a recorded diagnosis.[2] Thus, this isn't a study limitation as it would be in the US, so we have instead added this point to the Methods rather than the Discussion.	Methods, p. 5, lines 19–21

Comment	Response	Location of edits in track-changes revision
3) I am surprised the authors did not opt to compare the PAH patients to the “non-PAH” patients in terms of healthcare utilization and costs, particularly given that this algorithm for PAH patient selection is a different approach from other studies regarding healthcare utilization in the PAH population, and as such is not directly comparable to those studies. Additionally, the authors would have access to a substantial number of patients excluded from this cohort, and even simple descriptive statistics comparing included versus excluded patients would be informative and potentially enlightening. I would recommend the authors compare the PAH population to the non-PAH population to see if they are able to reach any conclusions regarding PAH patients as compared to non-PAH patients in terms of healthcare utilization	We agree that this comparison would have been valuable to perform, and indeed is something that should be done in future research. However, the present research was already challenging and adding this comparison would be a major undertaking. Given the complexity of our patient-ascertainment algorithm, we would not be able to develop the methodology to match non-PAH patients with PAH patients within the scope of the present manuscript.	–
4) Use of the “one year after start of dataset” index to separate incident and prevalent patients doesn’t make full sense. A patient could be newly diagnosed 2 months after start of the dataset, which would classify them as being “prevalent” by the criterion listed by the authors, even though this is a new diagnosis. This is further complicated by the classification algorithm used by the authors, which has multiple measures to classify disease, and as such is intended to identify “strict” PAH patients, but consequently is not as easy to use for something like incidence versus prevalence classification than a more binary tool like a date-associated ICD-10 code. I would remove this arbitrary classification scheme, and avoid referring to the “incident” versus “prevalent” patients in the results and analysis	We have added this as a potential limitation in the Methods. However, since PAH is a chronic condition, the number of patients who were classified as incident but were actually prevalent should be minimal. The issue is more that the “prevalent” patient subgroup could have included some incident patients. Thus, we have retained the separate analysis for incident patients, with the added caveat in the Methods. In addition, although previously some Results text and two table captions referred to prevalent patients, they actually present data for all patients, so do not in fact stratify patients by incident vs. prevalent. We have corrected this text accordingly.	Methods, p 4, lines 20–21 Results, p 7, lines 2, 4, 7 Table 1 caption, p 17 Suppl Table 6 caption, p 6
5) The authors note that the vast majority of outpatient visits lack an ICD-10 code to classify PAH-related from non-PAH related (Page 7, line 47-51). With that in mind, I do not see how these comparisons are made later in the manuscript (Table 2 describes outpatient visits stratified by PAH and non-PAH, but with only 10% of the data this would be misleading, a similar issue is seen in Table 4 and in Figure 5-B, Tables S2, S3, S6, and figures S-2 and S-3). Given the authors note this substantial limitation concerning outpatient visits in their data, I would not stratify the outpatient data into PAH and non-PAH.	We thank the reviewer for pointing this out. We have deleted this stratification for the outpatient results, adding an explanation in the Results text and revising the tables and figures in question accordingly.	Results, p 8, line 2 Table 2, p 18 Table 4, p 20 Figure 5, pp 25–27 Suppl Table 2, p 3 Suppl Table 3, p 4 Suppl Table 6, p 6 Suppl Figure 2, pp 8–10 Suppl Figure 3, pp 11–13

References

1. Sitbon O and Howard L. Management of pulmonary arterial hypertension in patients aged over 65 years. *Eur Heart J Suppl* 2019; 21: K29-K36. DOI: 10.1093/eurheartj/suz206.
2. NHS Digital. National Audit of Pulmonary Hypertension: Great Britain, 2018-19. Tenth Annual Report. 2019. (Accessed 1 November 2020, at https://digital.nhs.uk/data-and-information/publications/statistical/national-pulmonary-hypertension-audit/2019.)

3. Burke JP, Hunsche E, Regulier E, et al. Characterizing pulmonary hypertension-related hospitalization costs among Medicare Advantage or commercially insured patients with pulmonary arterial hypertension: a retrospective database study. *Am J Manag Care* 2015; 21: s47-58.
4. McLaughlin VV, Hoeper MM, Channick RN, et al. Pulmonary arterial hypertension-related morbidity is prognostic for mortality. *J Am Coll Cardiol* 2018; 71: 752-763. DOI: 10.1016/j.jacc.2017.12.010.