**EDITORIAL REQUIREMENTS**

REVISED MANUSCRIPT

The text of the revised manuscript should be double spaced and all pages should be numbered.

**REPLY: We addressed this accordingly.**

Two versions: (1) Submit your revised manuscript text file in Word, with revisions marked using the "track changes" or "edit" feature. (2) Also submit a clean version with all changes accepted.

**REPLY: We addressed this accordingly.**

TEXT WORD LENGTH

Limit the length of the text of your revised manuscript (not including title, abstract, references, and acknowledgment) to 3000 words.

**REPLY: We addressed this accordingly. However, given the plenty of modifications requested by the reviewers, the word count exceeds this by a little bit.**

SUPPLEMENTAL FILES

Please remove the supplemental file with tracked changes. Just the "clean" version is needed.

**REPLY: We addressed this accordingly.**

**EDITOR'S SPECIFIC COMMENTS:**

As noted by the statistical reviewer, it remains unclear the mechanism by which supplemental insurance would influence provider behavior. This in part relates to persistent confusion about whether the supplement/surplus payment is attributed per procedure (in which providers and hospitals may be incentivized to perform more procedures due to the related surplus) or per DRG (overall payment for the hospitalization, including procedures performed during the hospitalization, such that the incentive for providers in not directly evident).

**REPLY: Thank you for your continued interested in our work and further opportunity to clarify. Please also see our comment to the first question of Reviewer #1.**

**We have tried to better capture the peculiarities of the Swiss health system in a revised version of our eFigure 1 (see below and in the Supplement).**

**Generally, all providers receive a flat flee according to SwissDRG for every in-patient stay. This DRG payment is based on the financially, and not medically, leading cause of hospitalization. This provides a strong incentive for providers to perform profitable procedures in as many distinct/separate hospital admissions as possible and should be captured by our study design’s exclusion of emergency admissions.**

**We have rephrased the corresponding sentences in the introduction to capture this:**

***“For in-patients with basic insurance only, hospitals receive a standardized reimbursement as defined by the diagnoses-related group system (i.e., SwissDRG) which favors upcoding of diagnoses to the financially most lucrative cause of admission and leads to profitable procedures being performed in as many distinct admissions as possible. For patients with supplementary private insurance, hospitals receive the same flat fee as defined by SwissDRG plus an additional flat royalty which is individually negotiated between providers and insurers and economically more profitable. Additionally, physicians receive a sizable portion of this royalty on top of their regular salary further causing a non-medically driven incentive to fragmentize care into multiple admissions.”***

***(Page 5)***

Furthermore, there is question as to whether the observed association is a hospital or provider effect and the impact of clustering of patients with surplus insurance at specific hospitals needs to be considered.

**REPLY: We believe the effect is highly dependent on the provider, as both public and private hospitals care for both basically and privately insured patients (usually in a 2/3 to 1/3 ratio and vice versa) in dedicated wards (see updated eFigure 1 in the Supplement). Also, we have extensively addressed the issue of potential clustering in the reviewer’s comments and hope your concerns are properly addressed.**

**Reviewer #1:**

This paper investigates whether cardiac inpatients with supplementary private insurance are more likely to receive certain non-emergency intensive cardiac procedures if they are enrolled in supplementary private insurance. The authors find that, while the rate of intensive cardiac procedures has steadily increased across all coverage populations, those with the supplemental coverage were somewhat more likely to receive intensive cardiac procedures compared with those on basic insurance. This paper studies a very important question about how markets and insurance potentially misallocate care using a very clear identification based on a closed system with two similar but distinct coverage types.

**REPLY: Thank you for your kind words and suggestions.**

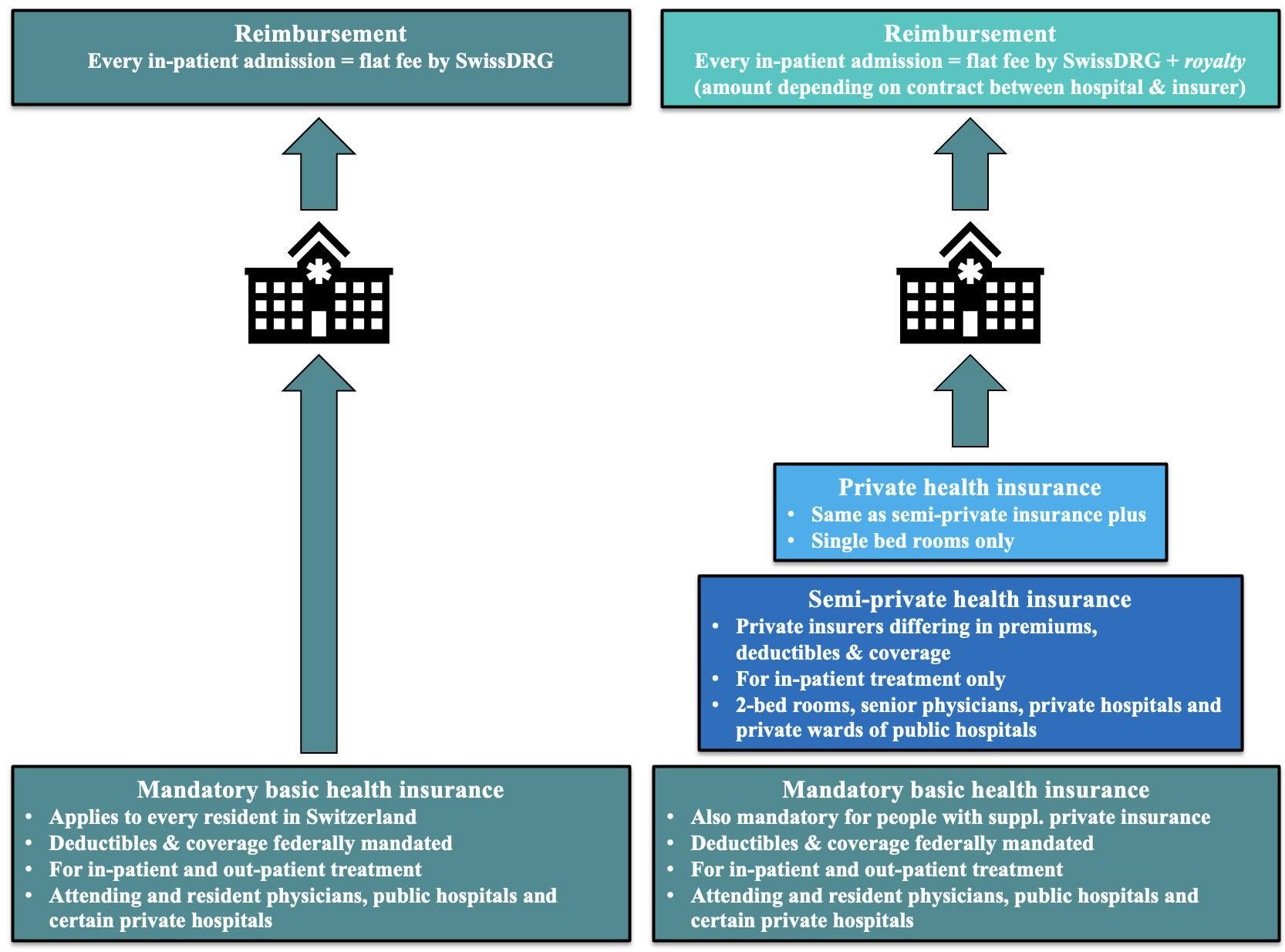
What is the difference between a surplus payment and a royalty? Are they the same payment for treating supplementary insured patients? They are both flat fees correct? Or does the size of the payment vary based on the DRGs?

**REPLY: There is no difference between surplus payment and royalty. We have changed the manuscript and now solely use the term “royalty” throughout. All providers receive a flat flee according to SwissDRG, plus royalties for both physicians and hospitals in case of patients with supplementary private insurance.**

**The amount of the royalty is not based on the DRG but rather on negotiations between each clinic/hospital and insurer. We have tried to better capture this in a revised version of our eFigure 1 (see below and in the Supplement). We have also rephrased the sentence in the introduction:**

***“For in-patients with basic insurance only, hospitals receive a standardized reimbursement as defined by the diagnoses-related group system (i.e., SwissDRG) which favors upcoding of diagnoses to the financially most lucrative cause of admission and leads to profitable procedures being performed in as many distinct admissions as possible. For patients with supplementary private insurance, hospitals receive the same flat fee as defined by SwissDRG plus an additional flat royalty which is individually negotiated between providers and insurers and economically more profitable. Additionally, physicians receive a sizable portion of this royalty on top of their regular salary further causing a non-medically driven incentive to fragmentize care into multiple admissions.”***

***(Page 5)***

****

My main question is why a patient with supplementary coverage would be more likely to receive additional or more intensive services if all care is fee-for-service. If the physician/hospital can only receive more payment in the form of a flat fee (surplus or royalty) for merely treating that patient regardless of the services rendered, this suggests a few possible explanations:

**REPLY: We hope your question was already answered by our response to your previous question. In essence, hospitals receive a fixed DRG fee for every admission, which provides a strong financial incentive to provide multiple procedures in as many distinct admissions as possible. Further, physicians receive royalties on top of their fixed salary, providing a strong incentive to perform additional, and potentially unnecessary procedures in distinct admissions. For further information on the Swiss health care system, we suggest the following reading (Individual Responsibility and Community Solidarity — The Swiss Health Care System, Nikola Biller-Andorno, and Thomas Zeltner, N Engl J Med 373;23 Dec 3, 2015, DOI:10.1056/NEJMp1508256)**

1.) A kind of exchange in which patients who desire these procedures trade the flat fees to access these procedures, even if their marginal benefit is low relative to the explicit cost (financial or resource cost) or the implicit cost (the value to their health of the next best and likely lower intensity cardiac intervention). This could be driven by patient selection into the private coverage (not great for modeling since this is the kind of selection trying to be addressed by the IPW approach) or though a kind of advertising/patient selection by physicians or hospitals.

**REPLY: Thank you for your remark. There are no precise data, but the available evidence suggests that most patients acquire their supplementary private coverage by their parents during childhood or subscribe to a plan in their 20ies and 30ies while they are still healthy. Once a person develops an ailment, it is merely impossible to obtain supplementary private insurance as insurers will decline offering a contract. We added this fact in the introduction section:**

***“Most people acquire their supplementary private coverage in their youth while they are still healthy 4. Once a person develops an ailment, it is merely impossible to obtain supplementary private insurance as insurers usually make extensive coverage exclusions or will not offer contract at all 4. This presumably leads to healthier people being more likely to get a supplementary private insurance based on a deceptive belief that more care is better, introducing a selection mechanism by which care allocation is inefficient.”***

**(Page 5)**

2.) Even if not overtly engaged in patient selection, private hospitals or senior physicians that are more likely to treat patients with the supplemental coverage may also tend to render more intensive cardiac procedures compared with what a public hospital would do for the same patient. If supplementary covered patients desire more senior physicians or private hospitals for other attributes, they may be unwittingly choosing providers who are more likely to treat with cardiac procedures.

**REPLY: We agree that there might be some hidden patient selection. However, we believe this effect to be rather small as hospitals still earn good money with procedures in basically insured patients. Additionally, the available evidence suggests that people with supplementary private insurance are healthier than similar patients with basic insurance only. We even assume that our effect estimates are conservative as the crude IR in our graphs consistently show a higher IR in patients with supplementary private insurance, despite their better health. Given that patients with supplementary private insurance are financially more lucrative, all public hospitals have dedicated private wards to compete with private hospitals. We have added this to the limitations section:**

***“One might be arguing that privately insured patients are unwittingly choosing providers who are more likely to treat them with procedures. Although there might be some patient selection in our model, we believe this effect to be rather small as hospitals still earn well with procedures in basically insured patients. Additionally, available evidence suggests that people with supplementary private insurance are healthier than similar patients with basic insurance only, which should lead to fewer procedures in privately insured persons 4. Hence, we assume that our effect estimates are conservative as the crude IR in our graphs consistently show a higher IR in patients with supplementary private insurance, despite their better health.”* (Page 14)**

**We tried to improve our analysis by further adjusting for potential clustering by hospitals using a mixed effects model (see below). However, results were not substantially different and model fitting was difficult why we decided to follow the law of parsimony and use the simpler model.**

These distinctions really matter not just from the standpoint of whether propensity score matching to address selection into insurance scrubs away some of these important dynamics, but whether the estimated differences in IRs is evidence of inefficient care use. If patients who are sicker appropriately overinsure, then the outcome would be more efficient. If patients with low risk overinsure based on a false belief that more care is better, then this selection is a mechanism by which the care allocation is inefficient.

**REPLY: As stated above, the Swiss health insurance system is in general providing/selling supplementary private plans only to young and healthy people. We therefore assume your second hypothesis to be the driving factor. To clarify this for other readers, we added this remark to the introduction section:**

***“This presumably leads to healthier people being more likely to get a supplementary private insurance based on a deceptive belief that more care is better, introducing a selection mechanism by which care allocation is inefficient.”***

**(Page 5)**

On line 155, the authors state that the primary outcome is the incidence rate ratio (IRR). This is somewhat confusing. In the negative binomial model, the outcome is some count variable, what seems to be the number of a certain procedure a unit of observation receives.

**REPLY: We agree that primary outcome is rather a rate than a ratio. Hence, we have rephrased that sentence and it now reads as follows:**

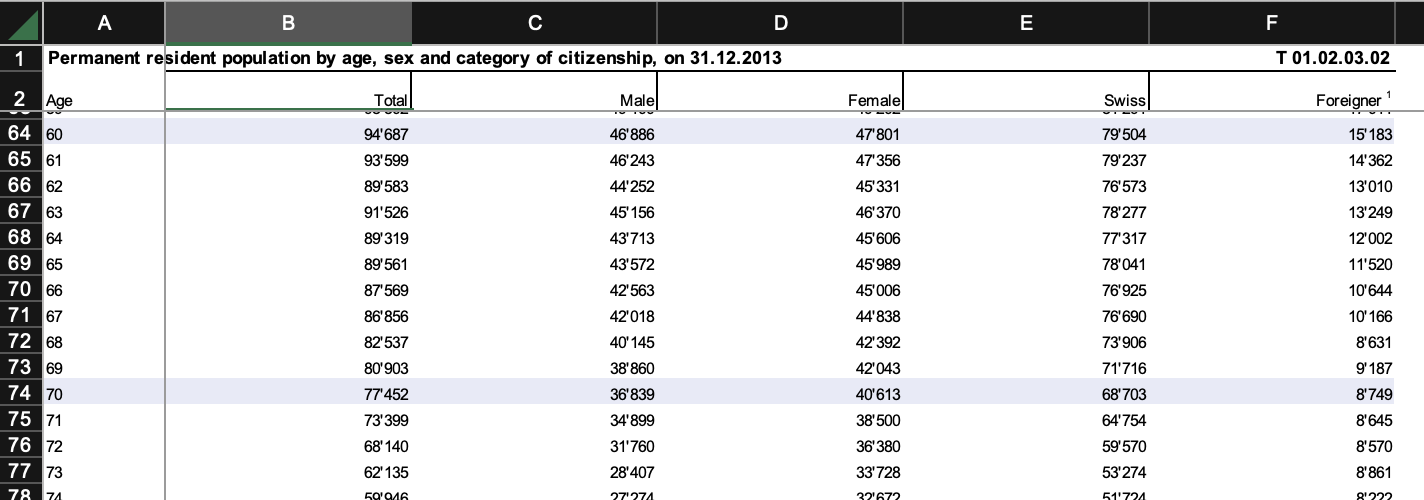
***“The primary outcome measure was the incidence rate (IR) of main cardiovascular procedures at primary position. Secondary outcome measures included IR of main and side cardiovascular procedures at any position, and secondary safety in-hospital outcomes were intensive care unit (ICU) stay, and ICU ventilation, 30-day readmission, and death.”***

**(Page 8)**

It seems the unit of observations is the aggregated unit of patients with same characteristics of year, age, and sex and “integrated data on person time.” What is person time in this context?

**REPLY: Because of the billing nature of our database, we did not have individual personal observation time to factor in our analysis.**

**As a work around, we used the number of people living in Switzerland which is publicly available and published as by the end of each year, sex, and nationality (Swiss citizen vs. aliens) by the Federal Statistical Office (see screenshot below for data snippet or** [**link**](https://www.bfs.admin.ch/bfs/de/home/statistiken/bevoelkerung/stand-entwicklung/alter-zivilstand-staatsangehoerigkeit.assetdetail.23064709.html) **for full data).**

****

**In detail, we counted the number of procedures for each year-, age-, and sex-specific group and paired the numbers with the published number of residents in Switzerland which belong to the same year-, age-, and sex-specific group. E.g., in 2013 there were 38,860 male persons at age 69 receiving 175 PTCA procedures in this group.**

**This leads to the following form of our dataset:**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Anonymous case number** | **Year** | **Female** | **Age** | **PTCA** | **Person time** | **PTCA counted across group** | **Propensity score** |
| 26888915 | 2013 | 0 | 69 | 0 | 38860 | 175 | 0.25382127 |
| 13590660 | 2013 | 0 | 69 | 0 | 38860 | 175 | 0.25034774 |
| 19706031 | 2013 | 0 | 69 | 0 | 38860 | 175 | 0.40772587 |
| 13296296 | 2013 | 0 | 69 | 1 | 38860 | 175 | 0.28194799 |
| 24304403 | 2013 | 0 | 69 | 0 | 38860 | 175 | 0.25275961 |
| 13100723 | 2013 | 0 | 69 | 0 | 38860 | 175 | 0.33324441 |
| 26267134 | 2013 | 0 | 69 | 0 | 38860 | 175 | 0.21684155 |
| 5110575 | 2013 | 0 | 69 | 0 | 38860 | 175 | 0.33508704 |
| 7377135 | 2013 | 0 | 69 | 0 | 38860 | 175 | 0.23637037 |
| 14484228 | 2013 | 1 | 69 | 0 | 42043 | 50 | 0.23707401 |
| 18166748 | 2013 | 1 | 69 | 0 | 42043 | 50 | 0.42652261 |
| 25720926 | 2013 | 1 | 69 | 0 | 42043 | 50 | 0.53390757 |
| 9677208 | 2013 | 1 | 69 | 1 | 42043 | 50 | 0.16286637 |
| 13780407 | 2013 | 1 | 69 | 0 | 42043 | 50 | 0.49203499 |

**We assumed one person-year for every resident, ignoring people dying or moving in and out of the country.**

**We revised and extend this part to the Methods section and provided a new subheader “Person time at risk”:**

***“Because of the billing nature of our data, we did have individual level data but no information on individual person-time to account for in our analysis. We used the number of people living in Switzerland at the end of each study year stratified by age and sex. This Swiss census data is publicly available and published by the FSO 12. We aggregated procedure counts on distinct strata per admission year, age, and sex. In detail, we counted the number of procedures for each of these distinct strata paired with the published number of persons in the same stratum, respectively. For instance, there were 38,860 male residents at age 69 receiving 175 PTCA procedures in 2013. We assumed one person-year for every resident, ignoring people dying or moving in and out of the country. We plotted IR curves after aggregating our dataset according to either age in deciles or monthly time steps and fitting a kernel-weighted local polynomial smoothing to illustrate changes over time. We calculated the difference of annual treatments in Switzerland based on the IR difference assuming a populace of 6.88 million (average population from 2012 to 2020, ≥18 years old) and 30% of people with supplementary private insurance 4.”***

**(Page 9)**

Also, how many of these aggregated groups are there in the negative binomial model? These are the true unit of analysis and will be significantly lower than the half million observations described.

**We aggregated across 9 years of observation, 2 sexes, and patients from age 18 to 105 (people older than 105 were collapsed into one bin by the Federal Statistical Office) resulting in 9\*2\*88 = 1,584 year-, age-, and sex-specific groups. However, we re-balanced by using the by number of observed procedures in the year-, age-, and sex-specific groups as the outcome and the same year-, age-, and sex-specific grouped number of residents in Switzerland as an offset (see below).**

**We agree that in theory we should lose some power with this approach. However, we do not consider it relevant considering our very tight 99% confidence intervals. We thought about splitting person-time further whether a resident is Swiss or a foreigner, however this would introduce another assumption since the FSO data cannot be stratified by sex and nationality simultaneously (compare with Excel spreadsheet screenshot from FSO data above). Furthermore, we think our approach is the only viable way to assess our research question (see your next question below).**

Why is this approach superior to an individual admission-level model in which each patient’s health or hospital stay data are controlled for individually (and the outcome would be the binary indicator of whether they received the procedure or not). While the propensity score matching and IPW approach do incorporate these characteristics, how are the individual-level propensity scores aggregated into each year/age/sex group? There will be considerable heterogeneity within each of these groups of patient comorbidity and other factors in the logistic model that guide clinical decision making about cardiac procedures.

**REPLY: We used the aggregated counted outcomes per year-, age-, and sex-specific groups as our outcomes but still performed the regression modelling on an individual level using the person time, i.e., the number of residents in Switzerland per year-, age-, and sex-specific groups as an offset:**

**E (PTCA count grouped per sex, year, and age) =**

**β0 + β1\*insurance\_class [IPW\_weight] + offset(person time grouped per sex, year, and age)**

**We have tried to fit a mixed-effects negative binomial model with the anonymous hospital identifier as a random intercept. However, in our dataset the number of observations is way larger the number of events leading to a numerical overflow in the statistical software (Stata v17) as the size of the log-likelihood number exceeds what can be stored by double precision (i.e., 64 bits).**

**As a work around, we have subsetted our dataset into 4 random parts and ran a regression on each of the four random subsamples. We also ran a sensitivity analysis with inclusion of the anonymous hospital identifier in the propensity score on the whole dataset:**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | **Mixed effects with random intercept model** | | | **Neg. binomial model (current)** | **Neg. binomial model with hospital ID in propensity score** |
| **Intervention** | **Subset** | **IRR (99% CI)** | | | **IRR (99% CI)** | **IRR (99% CI)** |
| PTCA main | 1 | 1.13 | 1.10 | 1.16 | 1.12 (1.12, 1.13) | 1.12 (1.12, 1.13) |
| PTCA main | 2 | 1.13 | 1.13 | 1.13 |
| PTCA main | 3 | 1.13 | 1.13 | 1.13 |
| PTCA main | 4 | 1.13 | 1.13 | 1.13 |
| LAA main | 1 | 1.20 | 1.13 | 1.27 | 1.15 (1.14, 1.16) | 1.15 (1.13, 1.16) |
| LAA main | 2 | 1.19 | 1.12 | 1.26 |
| LAA main | 3 | 1.18 | 1.18 | 1.18 |
| LAA main | 4 | 1.18 | 1.18 | 1.18 |
| MitraClip main | 1 | 1.18 | 1.09 | 1.27 | 1.08 (1.07, 1.10) | 1.08 (1.07, 1.10) |
| MitraClip main | 2 | 1.17 | 1.17 | 1.17 |
| MitraClip main | 3 | 1.15 | 1.15 | 1.15 |
| MitraClip main | 4 | 1.15 | 1.15 | 1.15 |
| TAVI main | 1 | 1.14 | 1.05 | 1.23 | 1.05 (1.03, 1.06) | 1.04 (1.03, 1.06) |
| TAVI main | 2 | 1.13 | 1.13 | 1.13 |
| TAVI main | 3 | 1.12 | 1.12 | 1.12 |
| TAVI main | 4 | 1.12 | 1.12 | 1.12 |
| PFO main | 1 | 1.00 | 0.99 | 1.02 | 1.01 (1.00, 1.01) | 1.01 (1.00, 1.02) |
| PFO main | 2 | 1.00 | 0.99 | 1.02 |
| PFO main | 3 | 1.00 | 0.99 | 1.02 |
| PFO main | 4 | 1.00 | 0.99 | 1.02 |
| Pacer main | 1 | 1.15 | 1.09 | 1.22 | 1.08 (1.08, 1.09) | 1.08 (1.07, 1.09) |
| Pacer main | 2 | 1.14 | 1.13 | 1.14 |
| Pacer main | 3 | 1.12 | 1.12 | 1.12 |
| Pacer main | 4 | 1.12 | 1.12 | 1.12 |
| A-fib main | 1 | 1.11 | 1.11 | 1.11 | 1.12 (1.11, 1.12) | 1.12 (1.11, 1.12) |
| A-fib main | 2 | 1.11 | 1.11 | 1.11 |
| A-fib main | 3 | 1.11 | 1.11 | 1.12 |
| A-fib main | 4 | 1.11 | 1.11 | 1.12 |
| Overall main | 1 | 1.13 | 1.13 | 1.13 | 1.11 (1.10, 1.11) | 1.11 (1.10, 1.11) |
| Overall main | 2 | 1.12 | 1.12 | 1.13 |
| Overall main | 3 | 1.12 | 1.12 | 1.12 |
| Overall main | 4 | 1.12 | 1.12 | 1.12 |

**The results from the mixed effects model are in our view too optimistic. Because of the minimal differences and problems in model fitting, we would prefer to use the most parsimonious model. Especially considering other potential options to improve model fit, e.g., rescaling of both outcome and person time variables, or using the negative binomial regressions log-likelihood values as input for the mixed effects negative binomial regression would introduce further, hard to understand bias. As such, we have added the anonymous hospital identifier variable to our propensity score model and re-ran all analyses and updated all results accordingly.**

**As we only have patients with an outcome/a hospital admission in our dataset, but no information on ambulatory care, we do not believe that a model using a binary indicator is appropriate. Given these limitations, the use of an aggregated count outcome together with aggregated person time is the most adequate solution to answer our research question. We acknowledge that this is a limitation, why we have added it to the limitation section:**

***“Furthermore, our database contains information on in-patient procedures only and to reduce bias, we have focused our study on procedures that are primarily performed in an in-patient setting. The restriction to in-patient procedures also made it necessary to make an assumption on person-time by aggregating it across admission year, age, and sex from publicly available data.”***

**(Page 14)**

The authors use propensity scores to facilitate an inverse probability weighting strategy to address potential endogeneity in selection into supplementary private insurance in which patient’s choice of supplementary coverage is potentially correlated with ultimate non-emergency cardiac treatment, possibly through some unobserved mediating factor such as proximity to a private hospital or desire to be treated by a certain physician. This, or the authors’ conception of why they need this multiple step procedure, could be made more clear in the paper paired with the kind of mechanism they believe they are capturing after adjusting for selection into supplementary insurance.

**REPLY: Although in some countries proximity to a hospital might be a leading factor, we do not think this is the case in Switzerland. Switzerland has way too many hospital beds, but national politics have failed to solve this issue despite numerous attempts over the past 2 decades. For instance, Switzerland has 46.5 hospital beds per 10,000 persons, while the United States has only 28.7 per 10,000 persons (both in 2017)** ([**https://www.who.int/data/gho/data/indicators/indicator-details/GHO/hospital-beds-(per-10-000-population)**](https://www.who.int/data/gho/data/indicators/indicator-details/GHO/hospital-beds-(per-10-000-population)) **).**

**We strongly believe that there is an inefficient risk selection by providers. Data shows that people with supplementary private insurance are better educated, have more disposable income, and are healthier than people with basic insurance only, while they are less frequently admitted to a hospital. Additionally, people spending more money for a certain good (i.e., health insurance) want to make sure they receive a better return (i.e., more care). Given these factors, we assume that people with supplementary private insurance seek out medical care more assertively, paired with financial incentives for providers this leads to an inefficient delivery of care.**

**Hence, we have added the following paragraph to the discussion section:**

***“Data from the FSO indicate that people with supplementary private insurance are better educated, have more disposable income, are healthier, and are less frequently admitted to a hospital than people with basic insurance only 4. It is also debated whether people spending more money for health insurance want to optimize their return by accessing more care. Considering these factors, we suppose that people with supplementary private insurance seek out medical care more vigorously. In turn, supplementary insured persons are more likely to meet providers delivering inefficient care backed by financial incentives (i.e., royalties) 13.”***

**(Page 12)**

Minor points:

In the introductions, it might be more clear to say ‘inefficient’ or outright ‘overuse’ of healthcare in lieu of 'inadequate’ which might be read as too little or insufficient care instead of too much care relative to the marginal benefit of that care.

**REPLY: In an earlier version of our manuscript, we also used the word “overuse”. However, we felt it to be a too strong word. Hence, we have now changed the text to your suggestion of “inefficient”.**

On line 138 under eligibility criteria, is it non-emergency procedure patients who are excluded or emergency procedure patients?

**REPLY: Thank you for catching this error. Of course, we wanted to state that we excluded emergency procedure patients.**

Can the authors identify information that might be associated with selection, such as whether the hospital is a private hospital, or any information about physician seniority?

**REPLY: Because we have billing data only, we do not have any information on physician seniority. Although, the database has an identifier for hospitals, it is pseudo-anonymized and access to the key is restricted to the Federal Statistical Office. We have added these facts to the limitations part of the manuscript:**

***“In addition, we were not able to adjust our model for patient characteristics not captured by diagnoses, such as income, education, living area, and laboratory parameters indicating socio-economic status, physician seniority, hospital ownership, and clinical severity.”***

**(Page 14)**